

1342

CACCTTGACT TGGGTGAGTA GGCACGGTAG GAGCTTTTCT CATAATCTCC TCTACCGTTG	360
ACAAGGAATC AGCCATGAGT TCTTCAGTTG AAGGTTCAAT TGCAGGAGTG CGAACTACTG	420
CCTCATCTTC TTTCAGAACT TCATCATAGC CTTTACTTTT TTCTAAATCT CTCAGAATCT	480
GCTCTTTAAA GCGTAATTTC TCTTCTGCTC TTGACTTTTC ACTCAAAAGT TTTTCCTCCT	540
TGTTGAGAAT CCATAATATT AGAGCTGAGA AGTCCAAAAA AAGCAATCTA TGATACTTTT	600
CCTAACGGAT TTTGTCATTT CCCAGACCAT ATCATACCAT GTTTCCTCTG CAAAGGTTGA	660
CTGGGAA	667

(2) INFORMATION FOR SEQ ID NO: 315:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1483 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 315:

GGGAAGCCAA GGTATTTTAT CGGATGAAGT TGTACTAGT TCTTCACCGA TGGCTACAAA	60
AGAGCTTCT AATGCAATTA CTAATGATTT AGATAATTCA CCAACTGTTA ATCAGAATCG	120
TTCTGCTGAA ATGATTGCCT CTAATTCAAC CACTAATGGT TTAGATAATT CGTTAAGTGT	180
TAATAGTATC AGCTCTAATG GTACTATTCG TTCCAATTCA CAATTAGACA ACAGAACAGT	240
TGAATCTACA GTAACATCTA CTAATGAAAA TAAGAGTTAT AAGGAAGATG TTATAAGTGA	300
CAGAATTATC AAAAAAGAAT TTGAAGATAC TGCTTTAAGT GTAAAAGATT ATGGTGCGGT	360
AGGTGATGGG ATTCATGATG ATCGACAAGC AATTCAAGAT GCAATAGATG CTGCAGCTCA	420
AGGGCTAGGT GGAGGAAATG TATATTTTCC TGAAGGAACT TATTTAGTAA AAGAAATTGT	480
TTTTTTAAAA AGTCATACAC ACTTAGAATT GAATGAGAAA GCTACAATTC TAAATGGTAT	540
AAATATTAAG AATCACCTTT CCATTGTTTT TATGACAGGT TTATTTACGG ATGATGGTGC	600
GCAAGTAGAA TGGGGCCCAA CAGAAGATAT TAGTTATTCT GGTGGTACGA TTGATATGAA	660
CGGTGCTTTG AATGAAGAAG GAACTAAAGC AAAAAATCTA CCACTTATAA ATTCTTCAGG	720
TGCATTTGCT ATTGGGAATT CAAATAACGT AACTATAAAA AATGTAACAT TCAAGGATAG	780
TTATCAAGGG CATGCTATTC AAATGTCAGG TTCGAAAAAT GTATTAGTTG ATAATTCTCG	840
TTTTCTTGGG CAAGCCTTAC CAAAACGAT GAAGGATGGG CAAATCATAA GTAAGGAGAG	900
CATTAGATT GAACATTAA CTAGAAAAGG TTTTCCTTAT GCCTTGAATG ATGATGGGAA	960
AAAATCTGAA AATGTGACTA TTCAAAATTC CTATTTTGGC AAAAGTGATA AATCTGGGGA	1020

1343

ATTAGTAACA GCAATTGGCA CACACTATCA AACATTGTCG ACACAGAACC CCTCTAATAT	1080
TAAAATTCAA AATAATCATT TTGATAACAT GATGTATGCA GGTGTACGTT TTACAGGATT	1140
CACTGATGTA TTAATCAAAG GAAATCGCTT TGATAAGAAA GTTAAAGGAG AGAGTGTACA	1200
TTATCGAGAA AGCGGAGCAG CTTTAGTAAA TGCTTATAGC TATAAAAACA CTAAAGACCT	1260
ATTAGATTTA AATAAACAGG TGGTTATCGC CGAAAATATA TTTAATATTG CCGATCCTAA	1320
AACAAAAGCG ATACGAGTTG CAAAAGATAG TGCAGAAaTWT TTAGGAAAAG TATCAGATAT	1380
TACTGTAAACA AAAAAATGTAA TTAATAATAA TTCTAAGGAA ACAGAACAAC CAAATATTGA	1440
ATTATTACGA GTTAGTGATA ATTTAGTAGT CTCAGAGAAT AGT	1483

(2) INFORMATION FOR SEQ ID NO: 316:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2453 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 316:

CCTGAACGCT TTTTATAAA TATCATAAAG CCAATCTGAT TTATCAAGTG TGTCTAAGCG	60
ACGCGAATTA AAATTCATTG CATACTCCAT CGCTTCTAAA AAACCTCATTT TTGAAAAGAC	120
GTAAAAATCA TCTAAATTCT GACTCCAATA TAATAACAAA ACCAATCCCA TAATATCCTC	180
TGTTTGATTA TTCAATAAAT TTAAGTTGGT TTCATAAAC CCTGGAGTTC CAAATAGAGG	240
CAACTTTTTT TCTTCAATTT GAGTTTCTTT CCTTAGGGCA TGCTCAAAGT CTATAATATA	300
AATATTATTT CTATTATCAA TAAGTATATT ATTAAATGAT AAATCTCTAT AGGAAAGATT	360
ATATTGGAG TTTATTATCT CCATATAATC AATTAATGTT AAAAACCAAT CATACGAGCC	420
ACTAACCATA TTATACTCGC TTAATTTATC TGCAATAATA AACTCAAATT CCACAAAATA	480
CGAATTCTTT ATGTAAAAAT CGTTAAAAAC TTTTGGAGTA AATTCCTCCT TTTCCAATTC	540
TACTAATATT TCTCTTCAT TTATTAAACG ATTCACAGAA TCTCTATTG TAAATCAAC	600
CAACGATAAA TCACTAGCTT CTTTAAATAA AGAATAAACT CGCTTTTGAG TATTAAATAC	660
TTTATAAACT CCACCTTTGG CATTTTTAGA AATCACTTCC AAAATAATAT ATTGATCAGG	720
AATAGTGTTA TATCTTGAA TATAGTAATC CCTTATTGGA ACATTACAT TTGAAGGGAT	780
TTCTTATCT CTTTATCCT TGAAAGTGCT ATCTTTTACG AACTCCCCAT ATCTGTAATA	840
TACAACCTCG CTAAGTGAA ATCTGAAATC TGATGGTATG TTTACACCCT TTACACCTTT	900

1344

ATACAATATT TCTAATTTGT GTAACAAACG TTGAAACTCT TTATTATCTT TTGGATAAAT	960
TGTAATGAAT TTCCCGACTT GTGAATAACC ATTAAGCCCT GTATTTTGCA AAGAAAGTTC	1020
TTTAATGCTA ACCAAAATTT TGAAATTTAT CTTCTTCTCT CTAGAAAATA TAAATCAAA	1080
GAATTTTTTA GCAACCAAAT TAGCATTAA TATTGAAGCG CTCAGGTGTA TTTTAAATCC	1140
CTTAGATTGG GTGATATTAG ACGGCAAATT ATATAACCAA TGTTCATCAC TAAATTTATC	1200
ACTAATTTTA TATTCTAATA ATAAATTATG GTATGCGTCT TCTATTTTCTAG TTTTCATAGTC	1260
CAATAGTTT AAATACTTTT CGTAATTCAT ATTAAGAAAT CTTCTCCATA AATTTTCTAGA	1320
CCATCATTTA AAGCCAAACA ATTTAAAGCG TGATAATAAA TGTGATAAT CAATGTAATC	1380
TTCAGTCCTC TATTTTGTA TTCCTTCACC AATAATTTTA TGCTATATCT ATTTTCTCGA	1440
GGCAATTTAT AGGACTTCAA GATAAAACCA TAAAGAGAT AAGTATTATA ATCTGACAAT	1500
CCAGTTTCAG AATAATTTT TAGAAAAATA TCTAGTGATT CTGATAATTC ATCCGGAATA	1560
ATTCTTTTAA CATCGTATTT ATTTTTCATA TCGGCCACTC TTCCTTAAAA AGCTCACAAT	1620
AAAATTTTAA ATTTCTATAC AACAATCCGA GAGTAGTCTC ACAATTTGAA CATTTTACAT	1680
CACTCTTAAT ATATAAAAA TGAATTAATC AGAAACCTCT GACTAAGATT TCCTAATTAA	1740
TTCACCTTCT ATATCATAGT AAGGAATTCT ATTATCCCTA ATTGAAAATT GAAATTTTAT	1800
GTTTTATATA TTAACAATTA TGCGGATTGT AAATCTTGTC TAACAAAATG GCAAGTGCTA	1860
CTATGTGCCC CAGAAGGCGA TGCAACGCTA TTTTGAATTG AAAGAGCATA ATCATCCATA	1920
TCATTTAAGT CACGGATTAG CAATGCTTCC TTCTCTCTC CGACAATTCC AAATTTTCTA	1980
ATTACCTTTT CAGGATTATC AAAAAATTCT CCAACAACCT CCATATTTCC TTGAAGTTCA	2040
TTCAAGAAAG CTTTCATTG ACTACTCATT ATATAGCTCC TTTTCTATTA CTTTATTTGG	2100
AATCAAACT TACTTGTAACA TTGGAAACAC CTCTATTCTA CGCTTTCATA TTGCTGCATG	2160
ACACTTTCAA AATCAAATTG CTAAAAATA TTTTAAAG CTTAATTTAG ATTTAATTAC	2220
ATATATCTCA AAAAATTGTT TTGAAATTAG TAAATTAAAA TAGGTTTCTG TACTTATAGG	2280
AAGTAGTTAT AAAAATTCG CCCATCATAA AATATCTATT TAAGTAAAC AAAAATTTTA	2340
TAATTTTTTG ATTTTAAAGT GACTATAATC TCCTATCTAT AAATACCATT CGCAGGACCT	2400
GGATCAATCC CTCTAGCCAT CTTATGAACT TGAGTTCCTC CAGACAGTCC CGG	2453

(2) INFORMATION FOR SEQ ID NO: 317:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1049 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

1345

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 317:

CCAATTTGAA GGCTCTAAAA CAATGGAAAA GTGCTACACA GATGTGACAG AATTTGCCAT	60
TCCAGCagTA CTCAAAAACT TTACTTTATCA CCAGTTTtag ATGGCTTTAA CAGCGAAATT	120
ATTGCTTTTA ATCTTTCTTG TTCGCCTAAT TTAGAATAAG TACAAACAAT GTTGGAACAG	180
GCATTCAAAG AGAAGCACTA TGAGAATACG ATTCTCCATA GTGACCAAGG CTGGCAATAC	240
CAACACGATT CTTATCATCG GTTCCTAGAG AGTAAGGGAA TTCAAGCATC CATGTCACGC	300
AAGGGCAACA GCCCAGACAA CGGCATGATG GAATCTTTCT TTGGCATTtT GAAATCGGAG	360
ATGTTTTATG GTTATGAGAA GAACTTTAGA TCTTTAGAAA ACCTTGAACA AGCTATTGTG	420
GACTACATTG ATTATTACAA CAACAAGAGA ATTAAGGTAA AGCTAAAAGG ACTTAGCCCT	480
GTGCAATACA GAACTAAATC CTTCGGATAA ATTAATTGTC TAACTTTtGG GGTGCAGTAC	540
ATTTTTGGTA TATATAAAAT TTGTAGGAGC TATATCTACA ATTTTATATT CCCAGTTTAT	600
GGATGTAACT TACTATATTC ACAATGTTAT CCAGTGTtTT TTCTCTAATA TTTAAGGAGT	660
GTTCTGTTTC TCGAATAAAT TCTTCAAAGT TTAACCCGTC AACTTGTtCC TGAACAAGAA	720
AATAATCATC CAGGATATAA AATTCATCAG TTAAATTAGT AGTATAACTT TTATCGGCTA	780
ATTTTTTTAG CATGTGAGCT TCATTTTTTA TATCATCAAG ACCTGTCCAT TCTCCTtCAG	840
CATCATAATT CACAAAAGGT CTTGACTGCT TGATGATTAC TTTTtTGCCCG TCCGATTtTC	900
TAATTGCCCG ATAAACATTt CCTTTATTG ATCTCTTAAT AATTTTTtCC ATTTTGTATT	960
TATTTATTGC AGAGTCCTTA CTTGAAACTT CACATGTGGT TTGAAAATAA ATCCTTTtTT	1020
CTTCTTCTGA AAATAAATCC ATTTTCCGG	1049

(2) INFORMATION FOR SEQ ID NO: 318:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 776 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 318:

TTAGTTGGTT AGAATCAGAA AATCGCCGAA GTGGTTATTT ATTTTtGAAT AAATTTAACG	60
AACCAATTAC AGCAAGAGGA GTTGCTCAAC AGTTAAAAAA TTATGCTGAT AAATACAAAA	120
TGAATCCTAA AGTAATTTAC CCTCATtCTT TTAGGCATTT ATTTGCTAAG AATTTTTTAG	180

1346

CGAAGTATAA TGATATTGCC TTGCTTGCAG ATTTGATGGG ACACGAAAGT ATAGAAACTA	240
CTCGAATTTA TCTAAGGAAA ACAGCTACTG AACACAAAA TATTGTAGAT AAAATTGTTA	300
ATTGGTAAAA AATAACAGGT GGTCAAACCTG ACTACCTGCT ATTTTGTGA TTATGGCTCT	360
TATTATGGGA ATATACCTAT GAATTGGGTT GTTATAAAAA TAAAAGATAT TTTTCAATA	420
AATACAGGTC TTTCTTACAA GAAGGGCGAT TTAAGCATTA ATAATAAAGG TGTTAGAATT	480
ATACGTGGTG GTAATATTAA GCCTTTAGAA TTTTCTCTGT TGGATAATGA TTAACACATT	540
GATACACAAT TCATCTCCTC TGAGCAAGTT TATTTAAAC ATAATCAGCT AATAACACCT	600
GTATCAACCT CTTTAGAACA TATTGGAAAG TTGCAAGAA TCGAGAAAGA CTATGATGGT	660
GTTGTGGCTG GTGGATGTAT TTTCCAATTA ACACCATTCG AAAGTGCAGA GATGATGTCA	720
AAATGTCTAT TATGTAACCT GTCCCTCTCCG TTATTTTATA AACAAATTGAA AGCAAT	776

(2) INFORMATION FOR SEQ ID NO: 319:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 658 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 319:

TGCAATGCGG CGGCTGCATA CGCTTGATCC GGCTACCTGC CCATTCGACC ACCAAGCGAA	60
ACATCGCATC GAGCGAGCAC GTACTCGGAT GGAAGCCGGT CTTGTGATC AGGATGATCT	120
GGACGAAGAG CATCAGGGGC TCGCGCCACC GAACTGTTTCG CCAGGCTCAA GGCGCGCATG	180
CCCGACGGCG AGGATCTCGT CGTGACCCAT GCGGATGCCT GCTTGCCGAA TATCATGGTG	240
GAAAAATGGCC GCTTTTCTGG ATTCATCGAC TGTGGCCGGC TGGGTGTGGC GGACCGCTAT	300
CAGGACATAG CGTTGGCTAC CCGTGATATT GCTGAAGAGC TTGGCGGCGA ATGGGCTGAC	360
CGCTTCCTCG TGCTTTACGG TATCGCGGCT CCCGATTCGC AGCGCATCGC CTTCTATCGC	420
CTTCTTGACG AGTTCTTCTG AGCGGGACTC TGGGGTTCGA TGTCGACAGC CCGCCTAATG	480
AGCGGGCTTT TTTTTCCTGA GGCTGGACGA CCTCGCGGAG TTCTACCGGC AGTGCAAATC	540
CGTCGGCATC CAGGAAACCA GCAGCGGCTA TCCGCGCATC CATGCCCCG AACTGCAGGA	600
GTGGGGAGGC ACGATGGCCG CTTTGGTCCC GGATCAATTC GCGCGACCGG ATCGATCC	658

(2) INFORMATION FOR SEQ ID NO: 320:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1475 base pairs
 - (B) TYPE: nucleic acid

1347

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 320:

CCGGCTTAAT TTTTAGAAAA CGTGGGCAGG GAACCTTTGT TCTCTCTCGT GGCAGCTCAA	60
AAAGAAAATT AATCGTTCCA GAAAGAGATA TCCGGGGACT GACAAAAATA TCTGAAGATG	120
CTCATTCTAC AATTGACTCG AGGATTATTC ACTTCAAATT AGAATTGCA AATGAATTTT	180
TAGCAGAAAA ACTACAGGTC GCTTTGCAGA GTCCAGTTTA TAATATTTAC CGCCTGCGTA	240
TTATTGACGG TAAACCTTAT GTTCTGGAAC AAACCTTATAT GAGTACCGAT GTTATTCCAG	300
GTATTACTGA AGATATTTTA CAAAAATCGA TTTACAATTA CATTGAAGGA AAGTTAGGAT	360
TGCATATTGC CAGTGCTACA AAAATCTTAC GAGCTTCTTC TAGTTCAGAA AATGAGCAAC	420
ATTACTTGCA GCTCCTTCCA ACGGAACCGG TATTTGAAGT AGAACAAGTG GCTTATTTGG	480
ATAACGGAAC TCCGTTTGAG TACTCGATTA GTCGTCATCG CTATGATTTA TTTGAATTTA	540
ATTCCTTTGC ATTACGACAT TCCTCCTAGG AGAAAATGTG AAAATGAAGC CAATCTTTTA	600
CAGACTCTAG TTTAAGAAAA ATTTAAACA GGGCAAGAAG GTCCCATCTA TGCTTAAATG	660
GTTTCTCTTT TCTAATAAAG ATGGCTTTAA AAGAGTGATC GTTGTATCCA TCATGTTGAA	720
AAATATCTTC GTATAGCTTA TAGACTAGGT ACTGAAATG TTCACCTGAT CTACTTCTTA	780
TAGTATTATTA GTTTTAAATA GTGTTTCAAA CATTCTTACA CTGACGAGAA GTTTTGTAGT	840
CTTTCTTGT AACACATATA GTATACTGTG GTTAGAATAG TAGACTGTGA CTTCTAACAA	900
ATTGCTAGAA ATGAATTTCA ATCTCCCAAT TTATTGTTC ATATCTTCTT TTAATATATT	960
AAATAAATTC TAAATCATAA TCATTTAAAA AAATTTTATT TTTTATTTT CATTACGAAT	1020
AATATAGATG AAGGGGAAAG AGTATGAAAA CAGAACTGTT TCTTTTGCTA TTAGTTCAAA	1080
AGGAGAAAAA ATGAAAGTAG AAAATATTTT GTATAGGGTG GATCATCGTA AATTGTTTGA	1140
TAATATTTCT TTTGATACTT CGAGTTCAGA CGTGACATTA ATTACTGGTA AAAATGGTAC	1200
AGGAAAGTCA ACTTTACTAT AGTAGATTGA AACTAGAATA GTACACATCT ACTTCTAAAA	1260
TATTGTTAGA AATCGATTG ACTATCCTGA TCTATTGTC CTGTTCTTAT TTCATTTTAC	1320
TATATCTCAA ATTGAGTATG ACGAAGTGCG CTCCCATGTC CTGGGAACGC ACTTTCTTCA	1380
TATTTTTCAT ATTCTTGAAT CCATCGATAA AGACTATTGG GATGAATTTT TAAAGTTGAA	1440
CTAATCATTT TTACAGGATG AGATTACAG CAGAG	1475

(2) INFORMATION FOR SEQ ID NO: 321:

1348

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 560 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 321:

GAAATATATA TACTTCATCT TAATAGTGAG CAAGCTAAAC TTAGCATTTT ATGCCCTCAT	60
ATGGGATGTT CTTTGACTAA ATAATATGAT TATCGAGATA TATCTGGATA AATGAACATA	120
TAAGTCTGAC GCGTAGACTT ATCAAAGTCA TTGGCATACA CCACTATGAA CTCGTTGGTC	180
TGTTCAAATC CCAACACATT ACCTGAGAAG AAAGTTGCAA TGTTGTTTTT GGTGCGGGTT	240
TGAATTTAAA AAATTTGTTA TGTAAGTACCT AATCTAAGGA ATTAGAACAA TGCCTCTAAT	300
TTTTCTTTAA TACACTGAAA CATTGATGAT TCTGGCTGTA TTTTGTAAAC AGCTCTTCTT	360
TGCTCCTGGA AAATATCTTC AGAAGTTATA TTCTCTATTC CTAACGCTAC TTGAGTTTTT	420
TTTCTAAAAA ATTCTTTTCC GTTGCCATCT TTAGAAAAAT CATAACCTTC CCTATCTACG	480
CTGTTACACA AATTAGCTAA AAAA r ACTCT GGGGTGGGA AAGGAAGATA AGAAaCGTAT	540
TTAGCCCATATA ATCTATAAAG	560

(2) INFORMATION FOR SEQ ID NO: 322:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 643 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 322:

CCGCCCCGCC ACCGCTGCCT ATCCTCGGGA GAGGGTCACC TGGAGTGAAC CTAGAACGAT	60
AGACACGGTG CGGTACGACC TCGTACTACT TTCGCCGACG GCCTCGTCCG TTGTCATCCA	120
CGAACTGATC GGACATGGGT GCGAACACTT CAGAGAAAAA ATCGTTGGAC TGCCTGTCGG	180
GCCTGAGGAA CTACGGGTGG TGGCTTTTCC GAAGAACGGC TCCGGGTTTG ATGACGAGGG	240
TACACCTCC GAAGAGATTG TACTTGTGGA GAACGGCATT GTGAGGCACG CTGTCAGGGA	300
TCGGGCGACT GGAGGAATGG CGCCTTTTTC CGGTTTGACC AAAGTGGCAT CACATGGTGT	360
CAACCTGGC TCAAGATGTA CGCATCTCAA GCGGAAGGG GAATCGTCAC AGGAAGGAGT	420
TACCGGAGTA CCCGCCGAAC GCACCGTTG GATAGAGCAT TTTTCTGCAG CGAACTACCA	480
TTCAGGTCGA GCCTTTTCA GGTCTGGCCT TGCTGGGTA GGCAGCCGAG AAGAATCTT	540

1349

ATATCCCTTA ATGCCTTTCA CCATGTCAAT TGATATCTAC GAACTGGCCA GCTTATTGTG 600
GCATTTAGAC GGTCAAACGG AACGAGCACG TAGGGTACTG TGC 643

(2) INFORMATION FOR SEQ ID NO: 323:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 780 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 323:

GGTACCCACT CATTCTTGAT GAATTGTGAA CAGTTGCCCT TGGGTCGTTT TGCGAGTTGA 60
AGTCAAGAAG AGGAAAAAAA CAAAAGGAG AAATACTCAT GGCAGTAATT TCAATGAAAC 120
AACTTCTTGA GGCTGGTGTA CACTTTGGTC ACCAACTCG TCGCTGGAAT CCTAAGATGG 180
CTAAGTACAT CTTTACTGAA CGTAACGGAA TCCACGTTAT CGACTTGCAA CAACTGTAA 240
AATACGCTGA CCAAGCATAC GACTTCATGC GTGATGCAGC AGCTAACGAT GCAGTTGTAT 300
TGTTCTGTTG TACTAAGAAA CAAGCAGCTG ATGCAGTTGC TGAAGAAGCA GTACGTTTCA 360
GTCAATACTT CATCAACCAC CGTTGGTTGG GTGGAAGTCT TACAACTGG GGAACAATCC 420
AAAAACGTAT CGCTCGTTTG AAAGAAATTA AACGTATGGA AGAAGATCGA ACTTTCGAAG 480
TTCTTCCTAA GAAAGAAGTT GCACTTCTTA ACAAACAACG TCGCGGTCTT GAAAAATPCT 540
TGGGCGGTAT CGAAGATATG CCTCGTATCC CAGATGTGAT GTACGTATG ACCCACATAA 600
AGAGCAAATC GCTGTTAAAG AAGCTAAAAA ATTGGAATC CCAGTTGTAG CGATGGTTGA 660
CACCAATACT GATCCAGATG ATATCGATGT AATCATCCCA GCTAACGATG ACGCTATCCG 720
TGCTGTTAAA TTGATCACAG CTAAATTGGC TGACGCTATT ATCGAAGGAC GTCAAGGTGT 780

(2) INFORMATION FOR SEQ ID NO: 324:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 624 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 324:

CGGGAAAAAT CAGATTGTGG GTTCAGATAT CGAATTAGCC AAGGCTATCG CAACAAAAC 60
AGGTGTCGAA TTGGAACAT CTCCCATGAG TTTTGATAAT GTACTGGCTA GTGTTCAATC 120

1350

AGGAAAAGCC GACCTTGCCA TATCAGGTGT TTCTAAGACA GATGAACGGA GCAAGGTGTT	180
TGACTTTTCC ATTCCTACT ATACTGCAAA AAATAAACTC ATTGTCAAAA AATCTGACTT	240
GACTACTTAT CAGTCTGTAA ACGACTTGGC GCAGAAAAAG GTTGGAGCGC AGAAAGGTTT	300
GATTCAAGAG ACGATGGCGA AAGATTGCT ACAAAATTCT TCCCTCGTAT CTCTGCCTAA	360
AAATGGGAAT TTAATCACAG ATTTAAATC AGGACAAGTG GATGCCGTTA TCTTTGAAGA	420
ACCTGTTTCC AAGGGATTG TGGAAAATAA TCCTGATTTA GCAATCGCAG ACCTCAATTT	480
TGAAAAAGAG CAAGATGATT CCTACGCGT AGCCATgAAA AAAGATAGCA AGAAATTGAA	540
AGAGGCAGTT CGATAAAACC ATTCAAAAGT TGAAGGAGTC TGGGGAATTA GACAACTCA	600
TTGAGGAAGC CTTATAAGCA TCCA	624

(2) INFORMATION FOR SEQ ID NO: 325:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1237 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 325:

TCTTATGAAG CCGAAGCGTG ATTTATGGCG GATAGGTTTG GTCTGCAGAA AGTGACAAAT	60
CTAGTGCCAT CAGCGTATAT GGAATCTnTG GCTGAGAAAC AGTCCCGGGG TGAAGTACT	120
TATGAGCAGG TTTATGAGGA TGCAACGGCT TATCATCATA CCATTGATGC GAGTACAGAG	180
GAGGCAGACT TGGTTTCTCT ACGTATTGTA GAACTATTGT CTCGAAGAGG CTTTAGCTTC	240
AGTCCTGCGA TCTTACTTGC TATTCATAAG GAGTTGTTTC AAGATATATT TGAACCTCG	300
ATTCGGTAG GTCAATTTTCG TCAGACTAAT ATCACAAGA ATGAACCTGT TTTGAATGGT	360
GAAAGTGTG TGTACTCTGA TTACTCCATG ATTCAAATGA CCTTGGATTA TGATTTTAAT	420
CAGGAAAAAC AAGTTGCATA TGCACACTA ACCCAGGCGG ATATGGTTAA AAAAATCCAG	480
CATTTTATTT CAGGAATCTG GCAGATTCAT CCATTTCGCG AAGGAAACAC TCGGACGGTA	540
ACGGTATTTT TGATTCAGTA TCTTCGTGAG TTTGGTTTTC ATATTGATAA TACACCATT	600
CAGCAACATT CCAAGTATTT TCGTGATGCC TTAGTGTTAG ATAATGCAAA GATTTTACAG	660
CGACGTCCTG AGTTTTTAAC AGCTTTTTTT GAAAATCTCT TGCTCGGTGG TCAAAATGAT	720
TTGTCTTCAG AAAAAATGTA TCTAGATTTA GACCTCGATC TTTCATAATC CTAATACTGA	780
GTAAACATTG AATTTTAGGA AAAAATGAAG TAAATATTCT CACAAGAAA CGTATATCAT	840
CAAAGTTTGG CTCTTTGTCA ATTGTAGTGG GTTGAAGAAA AGCTAAGTTC GAGAAAGGGC	900

1351

AAATTTCCGC CTTTCCTTTT TGATGTTTCA AGCGATAAAA ATCCGGTTTT TTGAAGTTTT	960
CAAAGTTTCG AAAACCAAAG GCATTGCGCT TGATAAGTTT GATGAGATTA TTGGGCGCTT	1020
CCAGTTTGGC ATTAGAATAG TGTAAGTTGAA GGGCGTTGAT AACCTTTTCT TTATCTTTGA	1080
GGAAGGGTTT AAAGACAGTC TGAAAAATAG GATGAACCTG CTTAAGATTG TCCTCGATAA	1140
GTTGCAAAAA TTTCTCCGGG TCCTTATTCT GAAAGTGAAA CAGCAAGAGT TTGAAGAGCC	1200
GATAGTGATG TATCAAGTCT TGTGAATAGC TCAAAAG	1237

(2) INFORMATION FOR SEQ ID NO: 326:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 461 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 326:

TTTGATTTTT CTGAATTAGA AGAGATTGAA TTGCCTGCAT CTCTAGAATA TATTGGAACA	60
AGTGCATTTT CTTTGTAGTCA AAAATTGAAA AAGCTAACCT TTTCTCAAG TTCAAAATTA	120
GAATTAATAT CACATGAGGC TTTTGCTAAT TTATCAAAT TAGAGAACT AACATTACCA	180
AAATCGGTTA AAACATTAGG AAGTAATCTA TTTAGACTCA CTACTAGCTT AAAACATGTT	240
GATGTTGAAG AAGGAAATGA ATCGTTGCC TCAGTTGATG GTGTTTGTG TTCAAAAGAT	300
AAAACCCAAT TAATTTATTA TCCAAGTCAA AAAAATGACG AAAGTTATAA AACGCCTAAG	360
GAGACAAAAG AACTTGCATC ATATTCGTTT AATAAAAAT CTTACTTGAA AAAACTCGAA	420
TTGAATGAAG GTTTAGAAAA AATCGGTACT TTTGCATTG C	461

(2) INFORMATION FOR SEQ ID NO: 327:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1436 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 327:

TAACATTTAG GTACCTCTTC TTAACAAAGT TCAATAGTAA CAATTAATAT TTAAACAAT	60
ATATCAAACA TCAATGACTA GAATACTTGC ATCATCCTTC TTTCCATAGA TTGGATCAAT	120
AGCAGAAGAA TTAAATCTCA TCTTAATTAA CTCTTCAAAA GTTTTATTTT GATTATTTTG	180

1352

ATAGAATTCA TAAAAGCCAT CGCTCATTA AACAATTTGT TCACTAGTAA CATCTATTTG	240
ATTAATAATA GCATGGTCTA AAAATCTCTC ATCCAACGAA CCTATCCAGT ACCCACTCGG	300
TTGATTAGAT AATTTTCTGA TTTTGTGTA AATAATTTT TTATTAAAA CACTATTTGT	360
ACCAATGAA TCTTTATCT CATTTTCCC TTTTCAAAT AAGTTATCTA CTCTATGATC	420
AGTTATTTCC ATTTCTGTTA CTAACATGAC GCAGTCACCT AGCATCATAT ACTCCAATT	480
TTTTCTGAA AGTTAGCAA ATATTGGTAA GCGATAATAT AGTATATTGA AACTAGAATA	540
GTACACCTCT ACTTCTAAAA CATTTGTAGA AATCGATTG ACTGTCCTGA TTGATTTGTC	600
CTATTATTAT TTCATTTTAC TATACTCTGT TAATTTATAT GAGTTTAAAC CGATTTTCATC	660
TTTAACCTCG AGTAAAGCAG TTTCAAATAT TTGTTTAAAG GTTTTGTATT CTTTACAATT	720
AACCGACAAA CTTTCTGATA AAATATGTAC AACTTCTGAG ACTGAATAAC CTATCTCCTC	780
TTTAGAATTA TATAAATCTG TAGCTCCACC AATAATCCAA AAATACTGAT TTGTGAACC	840
TACAATATCC TCATTTCTA CGGAACCTCC TTGTATCGAA CAAATTTTAT TTATCTTTAC	900
CATAACTCT CAACCTTTT AGTGTCAAAA GTAAACCAAT TCCTGTCACT GTTAAGAATA	960
GTCCATAAT CTTATTCGAA CCAGTCTTTG GTAATTTTGT TTTKACATCT ACTATyTCTT	1020
TAGATTTATT AATATGATTT TCAGTTCTC TGCCATCTCC AACTATTTTA TAGTTTACTT	1080
CTTCTGTCTT ATTATCTTGT TTATTGTCGA TCTTGTCATT CATTGTCTA TTATCTTTAC	1140
TTGAGTTAAA CTCTCCGTTT TCTGGTTAC TATCAATTAC ATTATTTGAA TTAGATTGTT	1200
TTTCTCTTT GTTTTTTCT TTTTCGTTT TATCACTTAA ATTATTTGTT ACAATTTTGT	1260
AAAGCCCATTT CTCCGTTACA ATATTGAAAT TACCATCGCT ATCACGTATA ACAGGTTCTT	1320
TCCCATTTGC ATTAGATTG ATGAATGATA TATACTTACC GGATAAATTA TAAATTTGGT	1380
TATTTAAAAC GGTATTTTA CCCTTGAAT CCTCAATAAC AATTCCTTCT TTACCC	1436

(2) INFORMATION FOR SEQ ID NO: 328:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 646 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 328:

CCGGCAGACA GGAGAAGGTG TTAAATATCA ATCTCAAATG GTTCGTCAAT GGTTTCTGAT	60
ACGTATTTTC CGTCTTCTT CCGTTGCTTG ACACACTCTG TGAGGAGATA TTCGATTGTC	120
CCATTGACTG AACGAAAGTC GTCTTCTGCC CATGATCCGA GTGCAGCGTA TAACTTTGTT	180

1353

GAGAGTCGAA GGGGGATCTG CTTTTTTTTA GCTTCAGCCA TCTTTAGTAA AGGCTTCCTG	240
TGTTGACAAT TGGTTGTGCA TCATGATTGC CACAAAGAAC GACAAGGAGA TTTGAAACCA	300
TGGCAGCTTT TCGTTCTTCG TCAAGTTCTA CCAATTCCCC TTCATTGAGC CGTTCTAGTG	360
CCATTTCAAC CATTCTTACA GCACCATCTA CAATCATCTT CCGTGATCA ATAATGGCAG	420
ATGCTTGTG GCGTTGAAGC ATAACGGCAG CAATTTCTGG AGCATAAGCT AGGTAAGTGA	480
TACGTGCTTC AAGGATTTC AAGCCAGCAT CCTCAACACG ACTTTGGATT TCTTCACGAA	540
TACGGGTAGC AACAAATTCG CTAGAGCCAC GGAGACTACC TTCATCTGCG TGCCCATCAC	600
CCGGAGTATC CACATTAGGA GACACATCGT AAGGATAGAT GCGGAC	646

(2) INFORMATION FOR SEQ ID NO: 329:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1653 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 329:

GTTGCAGGTG CAGTAGGTGT TACTTCAGAT ACATTTGAAC GTGCAGAGGC TCTTTTGTAG	60
GCAGGAGCGG ATGCGATTGT TATTGATACT GCACATGGTC ATTCTGCAGG TGTCTTGCCT	120
AAAATTGCCG AGATTCTGTC TCATTTCCCA GATCGGACTT TGATTGCTGG AAATATTGCT	180
ACTGCTGAAG GTGCACGTGC CCTTTATGAA GCGGGTGTAG ACGTTGTTAA GGTTGGTATT	240
GGACCAGGTT CTATCTGTAC TACTCGTGTG ATTGCTGGTG TTGGTGTTC GCAAGTAACA	300
GCTATCTACG ATGCTGCAGC TGTTCGCGGC GAATATGGTA AAACGATTAT TGCTGACGGT	360
GGGATCAAGT ATTCTGGAGA TATTGTAATA GCACTTGCTG CAGGTGGAAA TGCTGTTATG	420
CTTGGACTA TGTTTGCTGG AACTGATGAA GCTCCAGGCG AAACGAAAT CTCCAAGGA	480
CGTAAATTCA AGACTTACCG TGGTATGGGA TCAATTGCTG CTATGAAGAA AGGTTCAAGC	540
GACCGTTATT TCCAAGGTTT TGTCAATGAA GCAAACAAGC TTGTTCCAGA AGGAATTGAA	600
GGTCGTGTTG CTTATAAAGG AGCGGCAGCT GATATTGTTT TCCAAATGAT TGGTGGTATT	660
CGCTCTGGTA TGGGTTACTG TGGTGCAGCT AACCTTAAAG AACTACACGA TAATGCTCAA	720
TTTATTGAAA TGTCTGGTGC TGGTTTGAAG GAAAGCCATC CTCATGATGT GCAAATTACT	780
AATGAGGCAC CAAATTATTC TATGTAAAAA ACAATGAAAA GAACTCCAGT GAAAACAGGA	840
GTTCTTTTAC AATGTTGTCA ATTTCCATTT ACAGCAGCTT TACCATCCTG AATAGTGAAG	900

1354

ATACTTAGAT TTTCTGGCAG ATTTTGAAGA TGGTCTAAGC TTGTTGTTGT GATAAAGGTT	960
TGGATTGATT GAGAAATCGT TTCTAATAAT TTAACTGTC TAGTGTGTC AAGTTCACCT	1020
ATCACATCGT CAAGCAGTAA TATAGGAGAT TCTGTGGTAA TGCTTTCCAT TAATTCGATT	1080
TCTGCTAATT TTATCGAGAG GACGAGACTA CGATGTTGAC CTTGGCTTCC GAAACTAGCA	1140
TCCATCCCAT TTATATAAAA AGAAATGTCA TCTCGATGAG GACCGACACC AGTATTCTTT	1200
TTAAATAAAT CTCTGGATCT ACTTTTCTT AAAGCAATTT TGAAAGATTC GGATAAGTTT	1260
TGTTTGTCAG TTATATTGAC AGAAGATTGA TAGGATATTG ACAACTCTTC GATCTGATTA	1320
GAGAGTTCAA AATGTTTCTT ACGCCCAAAT GATTCTAGTT TTTTATGAA ATCTAAGCGG	1380
TGATTCATTA CACGACATCC ATAATCAACT AGCTGATCAT CTAACACAGA AAGGAATGTT	1440
TCATCTATTT TTTGAGCTGA TTTTAGGTAA GTGTTTCTTT GCTTTAGGAT GTGGTTATAA	1500
TTGGTTAAGT CAGATAAATA GATTGGCTTA ATTTGCCCAA GTTCCATATC AATGAATTTT	1560
CGTCGAATCG AAGGTGCTCC TTTAATTAGT TGTAATCTT CAGGAGCAA TAAGACAACA	1620
TTCATGTGTC CTACATAATC TGAAAGGCGT GCC	1653

(2) INFORMATION FOR SEQ ID NO: 330:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1340 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 330:

GAAACACTGT ATTTCAAAGC ATTTTGTGTT AGTTTAAAT TACTCCCAT CTCTTTTCC	60
AAACGTACAA TATATCCAAA ACCATTCAA ATACTAGATT CTATTTTTTA TAATATCACT	120
AAATCCACCT AATTATAGGA CGTTTTCAGA TTTTGTGTC CAGTCCAGT ACCGGAGAAA	180
TATTGTTTTA ATATAATATC TCTTTTGTG TTCTAAGCTC TTAAGCAA AAGAACAAGT	240
AAAGAGTCAA GACAAGGATA AAAAGTCCAT ATTAGGGCAA ATAAAAAGCT TTAAGACAGA	300
TGACAAATCT AAGTCAAATA AGAAAGACCA TAGCAAAGGT GCAGAGAGAT AAATATTGGC	360
GGTCTTCGGA CTGCCTTAT TTTTATATCC ATTTTCAA TCAAATTTAT TCAGACTATA	420
TATGCACATA TACACTTAAA TTCATATAA AACATGGCTT GTAAAAAATT ACTTTAATCA	480
CAATAATCGC ATTTAAATTT GTGATGTTG CAAGCTAAAT TACGGACTTC ACTTGAAGT	540
TTTCCCTTGT ATCTTTTATA ATAGATAGAA AATTGCTGG CAGATGAATA TCCAACAGAT	600
TCTGCTATCT CTTTATAGG TAGTTCAGTG TTAAAGAA GAGTTTCAGC TACATTCATT	660

1355

CTTTTCTTT GAGTGACTC TGTAATGCTT TGACAATATT TTTCCTTAAA TAAATTTT	720
AATTTAGTAC CACTCATT	780
AAATGATCAT CTAAGAATCT TGCTACATCT TCAAGTGCTT TATCATCATC AATTTCAATT	840
TTATATTTT TTCTATTAA GTATGTGTCA ATTACTATAC TTATCCATT	900
GCTTTAAAGA AAAAATCAGC GGCAGGAGCG TCCATCTTAC AATTTAATAT TTCCATTGCC	960
ACTCTTCTA AGGCCTTTGT AAGTATTATT TGATTGCTT GAAGCAAGGT TGAATAAAAA	1020
GATTCTGGAT TAATGTTAAT AGATGCTAAA TGTTTTCTA TTAGCTCTT TTTAAAAC	1080
ATGGAAACAG CAAGATAACA ACAATTCTCG TGTAATAAAA AAACAAAT ATCTTTTATA	1140
TTATCAAAAT CAAAGTACA TAGAGAGTTT GCGGTAATAG TTTGATACGG ATTAACTT	1200
TCTCCGTTG CACTGACAAT GAACTTGAA TAAATTGAAA CATAGTCTGA CATACTATAA	1260
GTGTATTTT GAACTACTT CTCTTGATA TAAAAATCAT GTATATCGAT AATGAAGATG	1320
CCTCCTTCAT AAAACCGGTA	1340

(2) INFORMATION FOR SEQ ID NO: 331:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 607 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 331:

TATGTCGTG ATGAGTTTTT AAGTAGGAAA AACGTGCTAA CCTCTCAGAT TTTGGAAC	60
GTAAAAGAAA CTCTTTTTT ACCCGTAGTA GTTGATAATG GGTGATCC GGCCTTAT	120
GAAATTGAGA AAAACAATT GCTAGCAAGT TTAGCAGCTG ATATGGATGA TTCTTTTAT	180
TTTGACATA AAGAATTGGA TAAATTGTTT TTTCATGATG AACGTCTCA ATTGGAATAT	240
AGTGATTAC GAAATCGTAT TTTAGCTGAA ACTCCACAAA GTTCTTATTC TTGTTCCAA	300
GAATTTTAG CCAATGATCG AATAGATTTC TTTTCCTAG GTGATTTTAA TGAGGTTGAA	360
ATTCAAAATG TATTAGAATC ATTTGGCTTT AAAGGTGCGA AAGGAGATGT GAAGGTCAG	420
TATTGTCAAC CTTATCTAA TATCCTCAG GAAGGTATGG TTCGAAAAA TGTGGACAA	480
TCCATTTTG AATTAGGTTA TCATTACTGT TCTAAATATG GTGATGAGCA ACATTTACCC	540
ATGGATTGAA TGAATGGTTT ACTTGGTGA TTTGCTCACT CTAAGCTCTT TACAAATGTC	600
CGGGAAA	607

1356

(2) INFORMATION FOR SEQ ID NO: 332:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 900 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 332:

TTAAATACC GAATTTTGTT TTGTCCTCTA TTTCAACATT GTGAATCGCC TCAGGCAGAG	60
AACCGATACT AAAGATATAA CCAAAATAGT TGTCATTGTC TTTACCGATA TCAATCTTAT	120
TGGTTAAATC AAAATCCAGT TCGTCAATTG CGCCATCGAT GTCTTGATTG ATTTCCAAAA	180
GTTTGTGAAT GAGGTTACCC GTACCGCCTG GGATAATCCC TAACTTAGGA ATGTAGTCTC	240
TCTCATCAAT ACCTGAAATG ACTTCATTGA CAGTTCCATC TCCACCAAAC ACAACCACTG	300
CATCATACTG CTCACGAGAA GCTTCTTCAG CAAAATGTGT TGCATCCAGC GCTTTTTCGG	360
TAATTTTGGT TTCAACATAT TCAAAGTATT CTTTGCCTTT ATTCTCCAGC TTTTCTTTGT	420
AATCCAAAGC CTTCTCGCCA CCAGAAGTAG GGTGATAAT TACCATTGCT TTTTTCATTG	480
ATTTTATCCT TAATTTTAAA CAGAAATGTT TACATTTTCGT CGTATGCAAG TAAATGTAAT	540
CCTATTATAC AATGAAAATA CAGAAAAGAG AAATCTGACG TACTGGAGAT TAATACGCTT	600
TTATTCTATT TTCCCATCGC CTAACATCAT CCTTTAAGGG TTCATCCAAG TAAGAATAGG	660
CCTTATCCTT GATCCAATCA GGAATACCGT AAGCTGCCTC TGCTAWGCTA CAAGTGATTG	720
CTGCGAGAGT ATCACTGTGC CCACCAAGTG AGATGGCATT TCTTATCGCA TCTTCGAAGT	780
CTCTACTTTC AAGAAAGGCG ATAATGGCTT GAGGGACAGT TTCCTGACAT GTTTCGTTAA	840
AACGATAGTT AGGACGGATT TCATCTAAAG TTTGAGATAG ATTGTAATCG TATTCTTTTT	900

(2) INFORMATION FOR SEQ ID NO: 333:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 533 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 333:

CCTTCTTGGC AACTGGTCT TGGAATACGG CAAAACCTCT GAAAATATCT ATGCTGGAAT	60
GGACGAGGAA TACCGTCGTT ATCAGCCTGC CATCATCACT TGGTACGAAA CAGCCAAACA	120
TGCTTTTGAT CGCGGACAGA TTGGCAAAAT ATGGGTGGAA TCGAAAACGA CCTCAAGGGC	180

1357

GGTCTCTACA GCTTTAAATC CAAGTTCAAT CCGACCATTG AGGAATTCGC TGGTGAGTTC	240
AACCTGCCAA CTAATCCTCT TTACCACCTC TCCAATCTGG CCTACACTCT CAGAAAGAAA	300
CTGCGCAGaA GcATTAAACAG AAAGGAAGCC TATGACCTTT AAACCTCTCA GCCAAGAAGA	360
ATTATCCAG CATACTCAG CTAGATCCCA ACGCTCTTTT ATGCAGACCG TAGAAATGGC	420
AGAGCTGCTG AGCAAGCGTG GCTTCAGTAC CCAGTATGTC GGCTACACTG ACCCACAAGG	480
GAAGGTAGTG GTGTCAGCTG TCCTCTACAG CATGCCTATG ACTGGTGGCC TTC	533

(2) INFORMATION FOR SEQ ID NO: 334:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 544 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 334:

CCAGCAAACT AGGAAGCTAG CCGTAGTTGC TCAAAGCACA GCTTTGAGGT TG TAGATAAG	60
ACTGACGAAG TCATGTACAA AACACTGTTT TGAGGTGCA GATAGAACTG ACGAAGTCAC	120
TCAAAACACT GTTTTGAGGT TGCAGATAGA ACTGACGAAG TCACTCAAAA CACTGTTTGT	180
AGGTTGCAGA TAGAACTGAC GAAGTCanna ACCACACCTA CCGCAAAGTG AATCTGAAGT	240
GGTTTGAAGA GAGTACAACT TGTCTTTTAG AAAAGGAGCC TATAATGAAA GTCTTTCAGC	300
ATGTAAATAT CGTGACTTGT GATCAAGATT TCCATGTTTA TCTTGATGGA ATCTTAGCAG	360
TCAAGGATTC TCAAATCGTC TATGTCGGTC AAGATAAGCC AGCGTTTTTA GAGCAAGCTG	420
AGCAGATTAT AGACTATCAG GGAGCTTGGA TTATGCCTGG TTTGGTCAAT TGTCACACCC	480
ATTCTGCAAT GACAGTCTG AGAGGGATCC GAGATGACAG CAATCTCCAT GAATGGCTCA	540
ATGA	544

(2) INFORMATION FOR SEQ ID NO: 335:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 349 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 335:

CCAGGAACTC AAATGTAAGT AGGGGTTCTT TTTTGTATA TTTTCAAAT AACGCCTCTA	60
---	----

1358

CACTATTTGT AGCAAATTCA CCAACTACAG TTGTATCTTA GTTAAAATAA GTTAGAATAT	120
GTAAGTGAGT ACCAGATATA CCAAGACATC GTCACCATCT AAGGTATATT CAAAATACAA	180
AAGTTGACCA ACTAGATTTC TGAATATCCT TATATATCCA TTCTTAAAT TGGTTTAAAT	240
AGCGTAGTCT TTAAACTAG TTTTGAGAAT CCAAAAAATC TTCCTACATA TGTAAGAAGA	300
TTTTTTAGTT CAGAATGATT AGATTAGCT AATGGATACC TATCCTACC	349

(2) INFORMATION FOR SEQ ID NO: 336:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1206 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 336:

CTCCGATAAC CACACCAGCA ATGGAAATAA TTCCATCGTT AGCATCAAGA ACACCCGCAC	60
GCAGGATATT TAAACGACCT GCAAAATTG AATCAATTTC GTGATTGTT TCTGACGCTA	120
AATTCAAGT TCAAGTTAGC CATCAAGAAG TCTTCTCTGG GTGACTTGTA GTCCAAGCAT	180
TTTTTAGGAT AGTTGTTAAT CCACTTTTCG ATGAATGCGA CTTCTTTGGG AGTCATTTTC	240
TTGGTTCCCT TAGGTAACCA TCTACGAATG AGCCTGTTGT GATTCTCATT AGTTCCCGGG	300
ATCCTCTAGA GTCGACCTGC AGGCATGCAA GCTTGGCACT GGCCGTCGTT TTACAACGTC	360
GATGACTGGG GAAAACCTG GCGTTACCCA ACTTAATCGC CTTGCAGCAC ATCCCCCTTT	420
CGCCAGCTGG CGTAATAGCG AAGAGGCCCG CACCGATCGC CCTTCCCAAC AGTTGCGCAG	480
CCTGAATGGC GAATGGGGCC TGATGCGGTA TTTTCTCCTT ACGCATCTGT GCGGTATTTT	540
ACACCGCATA TGGTGCACTC TCAGTACAAT CTGCTCTGAT GCCGCATAGT TAAGCCAGCC	600
CCGACACCCG CCAACACCCG CTGACGCGCC CTGACGGGCT TGTCTGCTCC CGGCATCCGC	660
TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGCATGTGT CAGAAGTTT CACCGTCA/C	720
ACCGAAACGC GCGAAACGAA AGGGCCTCGT GATACGCCTA TTTTATAGG TTAATGTCAT	780
GATAAGGATG GTTCTTAGA CGTCAAGTGG CACTTATCGG GGAAATGTGC GCCGAGACCC	840
TATTTGTTTA TTTGCTAAA TACATCAAA TATGTATCCG CTCGTGAGAA AATAAACCTG	900
ATAAATGCGT CAATAATATT GAAAAATGAA GAGTATGAGT ATTCTACATT TCCGTGTCGC	960
CCTTATACCC TTTTGTGCG CATGTTGCCT TCCTGTTTTT GCTCACCAG AAAACGCTGG	1020
TGAAAGTTTA AGATGCTGAA AAATCATTTG GGTGCACAAC TGGGGTTACA TCCAACGGA	1080
ATCTCCAnCA GCAGTTAAGA TCCTCTGACA GTTGATACG CCGCAAGAAC TATTCCCGAT	1140

1359

GAATGAGCAA CTTTAAAAAG TCCTGCGAAT GTTGGGGCGG TAATAATCCC CGTGTGTAG 1200
GCCCCG 1206

(2) INFORMATION FOR SEQ ID NO: 337:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 813 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 337:

CTGCTCAACT CAGACAGTCA AATTCTGAC TTTACCAAAA GAACCATCAA AAAAGTTGCT 60
GAAAAAGGCC ATCAGGTTAT TATTACGACA GGTGCGCCCTT ACCGTATGTC AAAAGATTTT 120
TACCGTGAAC TGGGCTTAGA CACTCCTATG ATTAACCTCA ACGGATCCCCT TACTCATTTA 180
CCAGACCAAG TTTGGGATTT TGAAGAGTGT TTGACTGTAG ACAAAAAATA TCTGCTAGAT 240
ATGGTTCAAC GTTCAGAGGA CATTCAAGCC GATTTTATCG CTGGAGAATA TCGTAAAAAA 300
TTCTACATTA CAAATCCCAA TGAAGAAATT GCCAATCCCA AACTATTTGG TGTAGAAGCT 360
TTCCAGCCTG AAGATCAATT CCAGCCTGAA TTGGTGACCA AGGACCCTAA CTGTATCCTC 420
TTCCAGACTA GAGCCAGTGA CAAATATTCCT TTGGCAAAAAG AAATGAACGC CTTCTACCAG 480
CATCAACTTT CTATCAATAC CTGGGGAGGT CCGCTCAATA TCCTTGAATG TACCCCAAAA 540
GGTGTCAACA AGGCCTTTGC TTTGGACTAC TTGCTCAAGA TAATGAATCG TGACAAAAAA 600
GATTTGATTG CCTTTGGAGA TGAACACAAT GATACCGAAA TGCTCGCTTT TGCTGGGAAG 660
GGTTATGCCA TGAAAAATGC CAATCCAGAG CTACTCCCTT ATGCAGATGA GCAAATTTCC 720
CTTACCAACG ACCAAGATGG GGTGCCAAA ACCCTACAAG ACTTATTCTT ATACCTATA 780
CTGATACTCA ATGAGGGGCA AAGAGCGAAC TTA 813

(2) INFORMATION FOR SEQ ID NO: 338:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 683 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 338:

CCTAGATAAA TGATATAATT CTATTATTGT TCGTAAAAAT TAAAAGGAGA TTGATGATGG 60

1360

ACAAATTATT TAAACTAAAA GAGAACGGTA CAGACGTTTCG TACAGAGGTT CTCGCTGGTT	120
TAACAACCTTT CTTTGCAATG AGCTATATTC TCTTTGTAAA CCCACAAATA CTTTCACAAA	180
CAGGAATGCC TGCTCAGGGC GTCTTCCTAG CGACGATTAT TGGTGCAGTA GCGGGTACCT	240
TGATGATGGC TTTTATGCT AACTTACCTT ATGCCCAAGC GCCAGGTATG GGAATCAATG	300
CCTTCTTTAC CTTTACAGTT GTATTCGGGC TTGGTTATTC TTGGCAAGAA GCCCTAGCTA	360
TGGTCTTCAT CTGTGGGATT ATTTCAATGA TTATTACCTT GACAAATGTT CGTAAAATGA	420
TCATTGAATC GATTCCCAAT GCTCTTCGCT CAGCTATTTT AGCTGGTATC GGTGTCTTCC	480
TTGCCTATGT AGGGATTAAAG AATGCTGGAC TTTTGAAATT CACGATTGAT CCAGGCAACT	540
ATACTGTTGT AGGAGAAGGG GCTGACAAAG CTCAAGCAAC GATTGCAGCA AACTCTTCAG	600
CAGTCCAGG ATTGGTCAGC TTTAATAATC CAGCTGTTTT AGTGGCTCTT GCAGGACTTG	660
CCATTACTAT CTTCTTTGTC ATC	683

(2) INFORMATION FOR SEQ ID NO: 339:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 852 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 339:

CTACTTTACA TGGAAGTAGT CACTGAATTC CAGTTAGAAA TTACTTTGTA ACTACGTTTT	60
GAGGAGGAGT AAAATGCTTT CCTACGTTTCG ATATTACCCA CTAGCGATAG CTAAATTAAT	120
GTGTCTGTGC TCTCCTAAAA TCTGCTGATT TATTACTGAC TAATACAGGA GGTTTTTTTT	180
ATGgACAGAC AATCATATCT GCTATTGCTG TTTATATTTC CACCAGTATC GATTATTTAA	240
TTATTTTAAT TATTTTATTT GCACAGCTAT CACAGAATAA ACAGAAATGG CATATTTATG	300
CGGGGCAATA TCTAGGCACA GGCTTACTTG TAGGGGCGAG TTTAGTTGCT GCTTATGTCG	360
TTAATTTTCG GCCTGAAGAA TGGATGGTTG GATTGCTTGG TTTAATCCCT ATCTATTTAG	420
GGATTCGCTT TGCAATTGTT GGAGAAGATG CGGAAGAAGA AGAGGAAGAA ATTATTGAAA	480
GATTAGAACA AAGCAAGGCA AATCAACTGT TTTGGACAGT TACATTGCTG ACAATTGCGT	540
CTGGCGGAGA TAATTTAGGT ATCTATATAC CTTATTTTGC TTCGTTAGAT TGGTCACAGA	600
CCCTCGTGGC CTTGCTTGTG TTTGTAATCG GCATAATTAT CTTTTCGAG ATTAGTCGGG	660
TGTTATCCTC TATTCCGTTA ATATTCGAGA CAATTGAAAA ATACGAGCGA ATCATTGTGC	720
CCTTAGTATT CATTCTACTT GGAATATACA TCATGTATGA AAATGGCACG ATAGAGACTT	780

1361

TTCTGATCGT GTAGATTTT TTGTTTCACT AGGGATTAG CCCGAGCTCA AATCAGCTCT 840
CTGATTTTCA GA 852

(2) INFORMATION FOR SEQ ID NO: 340:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 754 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 340:

CCGCACAAAA GCGCATAGTA TCAAGATCT ATAAAGCCTT GATACTATGC CTTTTTAATG 60
GATAAATAGT TAGTCTTTT TAAAGACCGG ATCTTTCAAA CTCTGCATAC TGGCATTGAT 120
CACCGCGCCT AGGATAACAA TTTTAGCAAT CAAGATAAAC CAAAACATCA TAACAACAAG 180
AAGAACGGAA CCTAAAATTC GGACATCCAC CAAATGATGG ACATAGTAAT TGAGATAACT 240
AGAGAACAGA GTTAGTAAAC CTAAATCAC TAAGAGAACA AAGGCACTGC CTGGTAGGGT 300
ATAGCTAATT TTCCTGTTAG ATAGATTGGG AAGAAAATAA TAAAGCATGA CCAAGATAGC 360
AAAGAGGAGG GCGTAAATCA GAGGACCTGC CAACCTTGT AAAGCCTGAT AGATAATGCC 420
ATCTTTTGTC CAATAATGAG CAAGTAAAGC CAAATCATC TGACCAATA AGATCAAAAA 480
CAAGGCAAA GCAAGAGGA GCTGCAACCA AACTGACTA GGAGACTTAG CATCTGATGG 540
GAAATAAGTC CACGACTCTT TTCGACGCCA TAAGCCTTGT TAAAGCTTT TTGCAAGAAA 600
TTCATAGATT TTGAAAACT CCATAACGCC GATAAACAG AAAAAGCTCA TAAACCTGTT 660
GAAGGTTGCG TCAAGACTTC TCTGGCTATT TTTCCACAC CTTCATAGAG GCTTGGGGGG 720
CAGACGTCTT TCATAAAGCC CAAAATTCT CCCA 754

(2) INFORMATION FOR SEQ ID NO: 341:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 707 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 341:

GGGGATAACT CTAGGAGTAC CGCTATTACT CGACTTAATG AGTGCACAAG AAGTCAGGAT 60
TTTTATGCAG GTTGGGCGCT TCATCAGACA GGGAAGATT ACAGCGACTA TTATGGAAGT 120

1362

CAAGGTTTGC TTTATTATTT GCTGACTTAC GTGAGTCAGG GCGGATTTTT CTTTGCCATC	180
TTTGAGTGGT TAGCCTTGGT AGCAGGAGGA TTTTTCCTTT TTAGATCAGC GGACACCTTG	240
ACAGAGCAAG GAGACCAAGC TGGACATCTG GTGACTATTT TTTACATGCT AGTTACAGGT	300
CTTGCTTTTG GTGGAGGCTA TGC GACTCTT TTAGCGCTTC CTTTCTTATT CGCAGCCTTT	360
AGTTTAGTTG CGGCTTACCT AAGCAATCCA AGCCATGATA AGGGATTTGT ACGGATTGGG	420
CTAGCTTTGG CAGGCGGATT TTTCTTTGCT CCCTTATCAT CGCTCCTGTT TATTGCTGTA	480
GTGAGTTTAG GCTTGTGGT CTTTAACCTT GGCATAGAC GCTTTGCGCA TGGGTTTAT	540
CAGTTTCTTG CAGTGGCTTT AGGTTTTTCA CTTGTCTTTT ATCCAACGTC CTACTATAGT	600
GCTGCAACAG GAAGTTTGG GGTGCGwTT AGTGGTATTC GTTATCCTAT TGACAGTATT	660
CGCTTTGATT TTA CTCTAA AATTTTAGAG AATATGTTTT TTAAAGG	707

(2) INFORMATION FOR SEQ ID NO: 342:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 762 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 342:

GGATTTTGAA AAACCATACC GATTTGACGA CGTATATTCC AAACATTTTC CTCAGTCAAA	60
CGTTGGCCAT CAATTACAAT CTCTCCGGAT TCTGCTTCCA GTAAGCCATC AATTAATCGA	120
ACCGTCGTTG ATTTACCACT ACCATTATGC CCTACAATCG AAAGCCATTC TCCACGTTTC	180
ACGTGAAAGT AATATCCTTC ACATCGTAGT AGTTCTGATT TTCTTTATAG CGAAAAGAAA	240
GATTTTTTAC ATCAATTATT GATTTCAATT CGAACCAAAT GTCCCTTTAA ATACATAGGC	300
ACTACCCTTG AAATAGTCAT AGCCAGAGTA GATAGTGAAA AATAAGGCTA CATAAAGTAG	360
AACTTGACCA AGCAAAGTCC AATGTAATAG CAAGAAAATA ATGGCAAACA TCTGACTAAA	420
AGTTTTAATT TTTCCAGGCA TTGCTGCTGC TAAAAATGTT CCACCAGTTT CAACCAATAA	480
AAGCCTTAAA CTTGTCACAG CTAACCTCACG ACAGATAATC ACTGCAACAA TCCAAGCCGG	540
AGCCATACCT AACTCAATCA ACATAATAAA AGCCGACATA ACTAGTAACT TATCCGCCAT	600
AGGATCTGCA AATTTACCAA AATTACTGAC CACATTCCAT TTACGAGCTA AATATCCATC	660
TAAATAGTCG GTAATACTGG CAACAGCAAA GATAATAGCT GCAACTATAT GACTCTCTAT	720
CGAATTTCTT ATCGTTAAAA TAAAGATAAA AATAGGTATA AA	762

(2) INFORMATION FOR SEQ ID NO: 343:

1363

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 482 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 343:

```
CTTTTGATAC ACTTAAACTA TGAATACAAA TCTCAAGCCC AAACCTCAGC GTTTTGCTTC      60
TGCGACTGCC TTTGCCTGTC CTATCTGTCA AGAAAATCTG ACTCTGTTAG AGACTAATTT      120
CAAGTGCTGC AACCGTCATT CTTTGGACTT GCGGAAATTT GGCTATGTCA ATCTAGTCCC      180
TCAATCAAG CAATCTGCTA ACTACGACAA GGAAAATTTT CAAAACCGTC AACAAATCCT      240
AGAAGCCGGC TTTTACCAAG CTATCTTAGA TGCTGTATCT GACTTGCTTG CAAGCTCAAA      300
AACTACCACA ACAATTTTGG ATATCGGTTG TGGTGAAGGA TTCTATTCTC GCAAACTACA      360
AGAAAGTCAC TCTGAAAAAA CTTTCTATGC CTTTGACATC TCCAAAGATT CAGTCCAAAT      420
CGCGGCTAAA AGTGAACCCA ACTGGGCAGT CAATTGGTTC GTTGGCGACT TGGCAGGACT      480
TC                                                                                   482
```

(2) INFORMATION FOR SEQ ID NO: 344:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 520 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 344:

```
TTTATTTTAA TAAAGTCAAT ACCTGTCTTT ACTTTTCTTT AAAAAAGTT TATTATGTTC      60
TTTAAGGAGG TGTA AACAT GAAAATAAAT AATAAACTCG TTGGAGAACG TATTCAAAAT      120
ATCCGTTTAA GCCATGGCGA CTCTATGGAA AAATTTGGAG AAAAATTAA TACTAGCAAA      180
GGTACAGTTA ACAACTGGGA AAAAGGTCGC AATTACCAA ATAAAGAAAA CCTACTAAAA      240
ATTGCATCTA TTGGAAAAAT GAGTGTGAA GAGTTACTCT ACGGCGATTA CAATACTTAT      300
CTACACTTAA AGATTATGGA TTTAGCTCCT GAATGTATAA AAAATTATGA TGAGTATAAC      360
TCTTTACAG ATGATATAAC AAATAAGCG TTACAGATCG CTCAAAATAC CATTTCTAAG      420
ATTGATTATC AAATTCAGA CGAAACGATC AAAAAATTTA TTGATTAGC TATCGAACAA      480
TCGAGAGATT TGCAAGGAAA TTTGTTGAAA AATAACGGGT                                     520
```

1364

(2) INFORMATION FOR SEQ ID NO: 345:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1003 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 345:

```
GCATCAAATC CGCCATCAAA GAAGTCTCTT GGATTACCA AGACCAGTCA AATAGCTTAG      60
AAGTGCTTAA TGACAAGTAC AATGTTCACT ACTGGAATGA CTGGGAAGTT GGAGACACGG      120
GAACCATTTG TGAGCGCTAT GGTGCCGTTG TTAAGAAACA CGACATTATC AATAAGCTTC      180
TCAAACAGTT GGAAACCAAT CCTTGGAAACC GCCGCAATAT TATTTCGCTC TGGGATTACC      240
AAGCTTTTCA AGAAACAGAT GGGCTGCTCC CGTGCGCCTT TCAGACCATG TTTGATGTTT      300
GGCGTGTTGA TGGGGAAATC TATCTGGATG CGACCTTGAC CCAGCGCTCC AATGATATGC      360
TGGTGGCCCA CCACATCAAC GCTATGCAGT ATGTGGCTTT GCAGATGATG ATTGCCAAAC      420
ATTTTGGCTG GAAGGTTGGG AAGTCTTCTT ACTTCATCAA CAACCTCCAT ATCTATGATA      480
ATCAATTTGA ACAAGCTCAG GAATTGCTCC GTCGGGAGCC GTCAAAGTGC CAACCACGCT      540
TGGTTTTTAA TGTTCCCTGAT GGGACTAATT TCTTTGATAT CAAAGCAGAA GATTTTGAGT      600
TGGTGGATTA TGACCCTGTT AAGCCACAGT TGAAGTTTGA CCTAGCTATT TAAAAGAATA      660
GAAAAAAGAA GTTGAGAATA ATCCCAACTT CTTTGTGTTT TTAACGTGAT ACGCGGCGAC      720
GAGCTGCTTT TTTACGGTTT TCTTCGATGA AAGCTGCTTT TTGCTCTTCT GGTTTCGATTA      780
CTTTCTTTTT AAATGCGTAT ACTGCACCTG CAACGGCAGC GACAGTTCCT GCGACACCTG      840
TTACAAGACC TTTAGCGAAT CCTTTAGCCA TGAGTCTTCC TCCTTTATAT TCTCAATCAG      900
CCAGCTCCTT CAAGAGGTCA CATTTTTCTG ACTGACCTTT TTGTGTTATA ATAATAGTAA      960
CGAAAAAATG GGAATTTTTC AAGGAAAAAA GATGAGAACA AAA                      1003
```

(2) INFORMATION FOR SEQ ID NO: 346:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 750 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 346:

```
CCGCACGTAC TATTCCAGAT GCCGAGGAAG TGGACCTCAT CCTCGTTGGC GCAACTGGTC      60
```

1365

TCAACGCCTT TGAACGCCTC TTGGTCGGCT CTTTCATCTGA ATACATACTC CGCCATGCTA	120
AGGTCGATTT GCTGGTTGTG AGAGAACAAG AAAAAACCTT ATAATCACAA AGAAAAGGAG	180
CCCCTAGCTC CTTTTTGTTC ACGATTATTT TCTCTCTTTA TGGCGTTCGT AAGCCTTGAG	240
CTGGCGCTGC AGTTCCTTTT TAATAGCAGG TTCTGGAGCA TATTTTTCTT CCCAATTATC	300
TGGTTTTAAG ATTTTATGGG TCACTGGATC AAAATGAGCC TTGCCATCTG GAAAAATTTT	360
CCCCATATTG GCCTGATGGA CAATATCAAA AATACGTTCT GGGTCCACCC CCATCAAGAC	420
AAAACTGCCG TAGGTGAAGT AAAGCGTGC AATCAAGGCA TCCACTTGCC CTATCAAATC	480
TTGCTGAGCA GGTGCTCTCT TGGCTACTTT ATCTGCTGCC TTATCAAGGG CCTGATGAAG	540
TTGCGATACA GCTTGACCAA AATCTTCTTC AGAAGGACTG GCTGCTCGAA CAAACTCCAC	600
CAATTCTTCT ATTTTAAAC CAGCCCTATG GGTGACCCC TCTAAATCCC AAGCTCGAGG	660
TTCTTCTTGG GTTCGTTTAT CCATCATGTG GTGGAAAGTC TTGACCTTAT TGAAATGATA	720
GTACCGGCTG ACAAGACTT TTTCTGAAGA	750

(2) INFORMATION FOR SEQ ID NO: 347:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 596 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 347:

CGCAACATAC GGATAACCTC CAAAGAATAT TTTTATATTA TAGCAAAGCT TTAAATTGAA	60
TGTTAGAGTC TTGTTCAAAA CAATCATCAA AACCACGTGG ATGATGGTAT TCTACTAAGT	120
GTTGATCTTG AGGATAAGTG TACTTACCGC CAACTTCCCA GATAAATGGA TGGAAATCGT	180
ATTGCAAGCG ATCTTTTCGC ATTTTCCAAA GTTCTAGAAT CTCATTAGTA GAAGCCATGA	240
AGTTAGACCA GATATCATAG TGAAGTGGGA TAATGACTTT GGTACGCAGA TTTTCTGCCA	300
TACGAAGAAG GTCGATAGAT GTCATTTTGT CTTGGATACC TACCGGATTT TCACCATAGT	360
TATTCAAAGC AACATCAATT TTAAAGTCTT TACCATGTTT TGCAAAATAG TTTGAGAAGT	420
GAGAATCTGC ACCATGATAG ATGGTTCCAC CTGGTGTTC AAAGATATAG TTAACAGCCT	480
TTTGAGCCAT TTCTTCATCT GTAACAGCCA AGCCAGCagT TCACCGCCTG TCTCATCAGC	540
ACCGTTCACT GGGAGAGTTA CCAAGCAAGT ACGGTCAAAT GATTCTACTG CATGAA	596

(2) INFORMATION FOR SEQ ID NO: 348:

1366

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 673 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 348:

CAGAGTCAAC AGCCTGAGTT GAAGGCAACT TTAGACACAG CAGTTACGAC AGCTGAATGA	60
GCTCCTCCAT CAGTTTTTTC TTTAATGAGT CCAGCTACAT CTTCAACTTC GAGGCCGTTA	120
ATCACAATGT CAGCGCCTAC TTCTTTTGCA AGGGCAAGTT TGTCAATTGT GATATCGACT	180
GCGATAACAT GAGCATTGAA TACTTTTTTA GCGTATTGAA CAGCGAGGTT ACCAAGTCCA	240
CCAGCACCCT AAAGAACAAC CCATTGGCCT GGTTCAACTT TTGCTTCITT GATAGCTTTA	300
TAGGTGTGTA CTCCAGCACA TGTGATAGAA GAAGCTTGGG CTGGATCAAG TCCGTCAGGA	360
ACTTTGACAG CATAGTCAGC AGTTACGATA CATTGTTTCA CCATACCACC GTCTACTGAG	420
TAGCCAGCAT TTTTCACTGT ACGGCAAAGG GTTTCGCGAC CAGTTGTACA GTATTGCAA	480
GTGCCATC CTTCAAAGAA CCAAGCAACG CTGACGCGGT CACCGACTTT AAGGCTTTTC	540
ACATCTGGAG CAATCTCTTT AACGATACCG ATACCTTCGT GCCCAAGAAC ACGTCCTGGG	600
ACTTGACCAA AGTCACCATG AGCAACGTGG AGGTCGGTGT GGCAAACGCC CACAGTATTC	660
ACTTCTACAA GTG	673

(2) INFORMATION FOR SEQ ID NO: 349:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 198 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 349:

GTACCCTACA AATGCTTTAC AGTATGGGTT GAGGGTGGTC AATGGAATA TGGAGTAGGT	60
TGGACAGGAA CTTTTCGATA TTCTGATTAC TTACATTCTA CTCGATATCA TACAGCAACT	120
GTTAGACATG GGGGTAGAAC CTCTAAGGAT TATGCAAAAC CTGAGGCATG GGCTAGAGCT	180
TCCCTCACCA AGATTCCG	198

(2) INFORMATION FOR SEQ ID NO: 350:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 891 base pairs
(B) TYPE: nucleic acid

1367

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 350:

GCTTCTTCTA TAGACAAAA TATCATGGGT AAAATAATCA AGGCTATAGC TAGAAGGAGG	60
GACCAATCCA CTACTAATCC TAAGAACAAA ACACTCAAGA GAGCAGAAGA GAGAGGTTC	120
CTGGCACTGA TAACGGCAAC CACCAAAGGA GAAACCAAGG ACACAGCCTT CATGGAAATG	180
AAAAAGCAA AAGCCGTTC AAAGAAAGCG ATAATGAGGC AAATCAAGAT ACTCCAAATA	240
TCAAGAGTAA AGGAAAGCTG ATAAACCGGC GAGAGGACAT TGCTAAACAA ACCTGCCAAA	300
ATCATCCCC ACCCAACCGT AGGAACAAAA CCATAACGCT TAGCAAAAGG TTGGGGCAAG	360
ATAACATTAA ACATAACACC CATGGCACTC AGCAAACCTG TTATAAGAGC TAGCGGCGTC	420
ATGGATAACT GAGAGAGGTC TCCCTTTGTC GCCATCAAGC AAACACCCAG CATGGCAACC	480
AAAACATAGA AAACAGCGCT TTTTGACGCT CGTTTTTGAT AAACCAAGCG ATTGTAAAAG	540
AGGATAAAGA CAGGGCTAAT AACTGTAAA ATAGTTGCTG TCGTAGCATT TGAGTATTCT	600
ACACAGAGAT AGAAAAATA CTGAACTGAA AAAATCCCCA AAATAGCATA GGCTAAAAAG	660
GGCAGGTAAT TTTTCTTGTC TCGCCAAATA TCTAGCACTT GCGATTTTAA TTGTATTGCA	720
GACCAATGA GTACAAGACT CCCTGCCAGT GTCAACGCA TACAGCTAAT CCAGCCCGAA	780
GACACCTGAT AATGAGTAAA GAAGTACTCT CCTAAAATTC CACAGATTCC CCATATTAAG	840
CCGATAGGA GCGAATAAAT TTTCCGTTA ACAATCTTTT TCTGATACTG A	891

(2) INFORMATION FOR SEQ ID NO: 351:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 325 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 351:

GAAAGCGTTC AATAGAACAT TGCTTTTTTA TTTTtagagt AAGCTAAGCG CTTCAGCATC	60
TGCGATGATG GTTACATCAG GGTGATTTTG GAGGCTACTT GCAGGTAGGT TCTCAGTCAC	120
TGGGCCAGAT ACTGTTCCGG CAATGGCTTC TGCTTTCGAC TCACCGTAAG CAAAAAGAAT	180
AATAGACTTG GCATCCAAAA TGTTTTTAAT CCCATTGAA ATAGCTTGGG TTGGGACGTC	240
TTCAATCTTG GCAAAGAAGC GTGCATTGGC TTCGATAGTA GACTGGTCAA GTTCTACTAG	300

ATGCGTTTGA CTGTCAAATG GAGTG 1368 325

(2) INFORMATION FOR SEQ ID NO: 352:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 344 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 352:

CAAGAGCAGT TTGATGATTT TTGATAAGCA TGCGAATTTA AAATACAAAT ATGGCAATCG 60
CAAGTTTGG TGTAGAGGCT ATTATGTAGA TACGGTAGGC CGTAATCAGA AAGTGATAGC 120
TGAATATATT CAGAATCAAT TACAAGAAGA CAGAGTAGCA GACCTAGCTC ACGTATTTCG 180
AGTCAGTAGA TCCGTTTACT GCGGAAATAA ATAAGAGGAA GTAACGTnAA GTGCTTTAGC 240
ACCTGCTCGG GAAAGTGGTG CCGAGGAAG CTATTTTCAGG ATGCTTTGGC CCTGGCCGGT 300
AGAAGCGTTA TAGCCGCGA CTACGACACT TCACACTGGT GGTT 344

(2) INFORMATION FOR SEQ ID NO: 353:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 692 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 353:

CCCTATCCCT GCTATTGGGG CTGCTCTCAT TGCTGCTTTG GCACAAATCA GTCTTCCAAT 60
TGGACCTGTT CCCTTCACTC TGCAAACTT TGCAATCGGC TTGATTCTAC TGTCTTTAGA 120
CCGAGAGAGG CTGTACTTTC TGCTGGACTC TATCTTCTTC TAGGTGCTAT CGGTCTTCCT 180
GTCTTTGCAG GAGGTGGAGC TGGTTTTCAG GCTTTAGTTG GCCCTACTGC AGGCTATCTT 240
TGGTTTTATC TCGTTTACTC TGGACTTACT TCCTCTCTAA CCAACAGCAA GAGTGGTGTT 300
GTTAAGATTT TTCTTGCAA CCTCTTGGGT GATGCCCTTG TCTTTGTCGG CGGGATTCTC 360
AGCTTGCAAT TCCTAGCTGG AATGGCATTG GAAAAGCTC TTGCTGTGGG GGTCTTCCC 420
TTTATCATTC CAGACCTGG CAAACTTCTA GCTATTAGTT TTATTAGCCG TCCCCTACTT 480
CAACGCCTTA AAAATCAGGC TTACTTTACT AACTAAAAA GGATATCGAG TTATCATGAC 540
TCAATATCCT TTCTTTTAT TTTGAAACT TATACTCAAT GAAAATCAA GAGCAAATA 600
GGAAGCTAGC CGCAGGCTnG CAAAACACTG TTTTGAGGTT GTGGATGAAA CTGACGAGTA 660

1369

AnATCTCATA CATACGGCAA GGCAAAGCTG AC

692

(2) INFORMATION FOR SEQ ID NO: 354:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1005 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 354:

GTGATGGACT ACTGGTTCAA AACGCATCCA GAAGATTTTT TCGATAATGT CGGACCTCTT	60
GTAGCCAGTA ACTTTTTTCA TACTTACACC GAAGATTTCC ACTTGATGAA GGAAATTGGA	120
GTTAATTCTT TCCGCACTTC CATCCAATGG AGTCGACTCA TCAAGAATTT AGAGACAGGT	180
GAGCCTGATC CAAAAGGTAT TGCTTTCTAC AATGCCATCA TTGAAGAAGC TAAAAGAAGC	240
CAGATGGATC TTGTGATGAA TTTACATCAT TTTGATTAC CAGTGGAAC TCTTCAAAAA	300
TACGGTGGTT GGGAAAGCAA ACATGTAGTG GAGTTATTCG TGAAGTTTGC CAAGACTGCT	360
TTACATGCT TTGGAGATAA GGTTCATTAC TGGACAACCT TCAATGAGCC AATGGTCATT	420
CCAGAAGCAG GGTACTTATA TGCTTTCCAT TATCCAAATC TAAAAGGAAA GGGAAAAGAG	480
GCCGTACAAG TCATCTATAA TCTAAACCTT GCTAGTGCAA AAGTGATTCA ACTATATCCC	540
TCATTAGAAC TTGATGGAAA GATTGGGATT ATTTTAAACT TGACACCTGC TTATCCAAGA	600
AGTAATTCTC CAGAAGACTT AGAAGCAAGT CGATTTACAG ATGACTTCTT TAACAAAGTC	660
TTCTTGAATC CAGCTGTTAA AGGAACTTTC CCAGAAAGAT TGGTAAAACA GCTAGAGAGA	720
GATGGCGTGT TATGGAGTCA TACCGAAAAA GAGCTTCAAC TGATGAAATC AAATACGGTT	780
GATTTTCTTG GAGTAAACTA CTACCATCCA AAACGTGTTT AAGCACAGC AAATCCTGAG	840
GAATATCAGA CGCCCTGGAT GCCAGACCAA TACTTCAAAG AGTATGAATG GCTGGAGCGT	900
CGCATGAATC CATATCGTGG TTGGGAAATT TTTCCGAAAG CCATTTATGA TATTGCTATG	960
ATTGTGAAGG AAGAATATGG TAATATCCCA TGGTTTATCA GTGAA	1005

(2) INFORMATION FOR SEQ ID NO: 355:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 973 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

1370

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 355:

CCGACAAGCA ATATTAAAAA GAGTAAACTA TTAAGTAGTT AATTAACCGG TTTATTACTT	60
TATAGTGAAT CAAATATACT TAAGAAAAGA GGAAAGAATG AAAATTAATA AAAAATATCT	120
AGCAGGTTC A GTGGCAGTCC TTGCCCTAAG TGT TTGTCC TATGAGCTTG GACGTTACCA	180
AGCTGGTCAG GATAAGAAAG AGTCTAATCG AGTTGCTTAT ATAGATGGTG ATCAGGCTGG	240
TCAAAAGGCA GAAAACTTGA CACCAGATGA AGTCAGTAAG AGGGAGGGGA TCAACGCCGA	300
ACAAATTGTT ATCAAGATTA CGGATCAAGG TTATGTGACC TCTCATGGAG ACCATTATCA	360
TTACTATAAT GGCAAGGTTC CTTATGATGC CATCATCAGT GAAGAGCTCC TCATGAAAGA	420
TCCGAATTAT CAGTTGAAGG ATTCAGACAT TGTCAATGAA ATCAAGGGTG GTTATGTCAT	480
TAAGGTAAAC GGTAATACT ATGTTTACCT TAAGGATGCA GTCATGCGG ATAATATTCC	540
GACAAAAGAA GAGATTAAAC GTCAGAAGCA GGAACGCAGT CATAATCATA ACTCAAGAGC	600
AGATAATGCT GTTGCTGCAG CCAGAGCCCA AGGACGTTAT ACAACGGATG ATGGGTATAT	660
CTTCAATGCA TCTGATATCA TTGAGGACAC GGGTGATGCT TATATCGTTC CTCACGGCGA	720
CCATTACCAT TACATTCCCTA AGAATGAGTT ATCAGCTAGC GAGTTAGCTG CTGCAGAAGC	780
CTATTGGAAT GGAAGCAGG GATCTCGTCC TTCTTCAAGT TCTAGTTATA ATGCAAATCC	840
AGCTCAACCA AGATTGTCAG AGAACCACAA TCTGACTGTC ACTCCAACCT ATCATCAAAA	900
TCAAGGGGGA AACATTTCAA GCCTTTTACG TGAATTGTAT GCTAACCCCT ATCAGAACGC	960
CATGTGGGAT CTG	973

(2) INFORMATION FOR SEQ ID NO: 356:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 843 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 356:

GGTCGCATCT GCAATATCTG TCGCCTCCAC ATAAGCGACA CCAGCCTTGT CTGCTGCCCCG	60
TTTGACACGT TCTGCAGATT GACCCAGGAT GACCATCTTC TTGAGTCCAG TAATGTCTGG	120
CACCAATTCTG TCAAACTCAT TGCCACGGTC CAAACCACCT GCAATCAAGA CGACCTTGCT	180
GTTGTCAAAT CCTGACAAGC TTTTGTAGTA GCCAAGATAT TAGTTGATTT ACTGTCGTTA	240
TAGAATTTAA CACsCTTGAT GTCATCCACA AACTGGAGAC GGTGTTTGAC ACCACCGAAG	300
GCTGAAAGAG TTTCTTGAT GGT TTGATTG TCCACATCAC GAAGCTTGCC TACAGCAATA	360

WO 98/18931

1371

GTCGCAAGGG CATTITCCAC ATTGTGGCTA CCTGGAACAC CGATTTCATT CGCTGCCATG 420
 ACTACTTCAC CACGGAAGTA GAGTTGACCA TCTTCCAGAT AAGCTCCATC AACCTTTTCA 480
 AGTGTTGAAA ATGGTACAAC AGTGGCTTCT GTCTTGAAG TCAAGTCTTT TGCCAAGTCT 540
 TGATTAAAGT TCAAGACAAG GAAATCAGCT GCTGTCATCT TGTCTGGAT ATTCCACTTG 600
 GCTGCTACAT ATTCCGAAAA TGACCCATGG TAGTCGATAT GAGTTGGCAT GAGTTGGTA 660
 ATAACCGCAA TCTCTGGATG GAATCTTGA ACACCCATGA GTTGAAAGA AGAAAGTTCC 720
 ATAACAAGCG TGCCTTATC TGATGCTATT TGAGCAACCT GACTAGCTGG ATAGCCGATA 780
 TTCCCTGATA AAAGACCATG TTGGCCAGCA GCAGTCAAAA CTTCCCGGGT TCCTCTAGAG 840
 TCG 843

(2) INFORMATION FOR SEQ ID NO: 357:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 807 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 357:

TTTTTTTTAT ATTTTTTTTA TTTATTATT TTTGGCAAAA AAGACCAATT TGCTTTGGAG 60
 CATTTGCTTCT GCATTAAAT GTCTATTTT GCTCGTGCTG TTACGCTCTT TGTATCATGT 120
 ATTAAGTAGC AAGTGCAACT TGCAAACTAC TAGTAAGAGG AGAAAAACAA AATGGTTATG 180
 ACTGACCCAA TCGCAGACTT CCTAACTCGT ATTCGTAATG CTAACCAAGC TAAACACGAA 240
 GTACTTGAAG TACCTGCATC AAACATCAAA AAAGGGATTG CTGAAATCCT TAAACGCGAA 300
 GGTTTTGTAA AAAACGTTGA AATCATTGAA GATGACAAAC AAGGCGTCAT CCGTGTATTT 360
 CTAAATACG GACCAAAATGG TGAGAAAGTT ATCACTAACT TGAACGTGT TTCTAAACCA 420
 GGACTTCGTG TCTACAAAAA ACGTGAAGAC CTTCCAAAAG TTCTTAACGG ACTTGGAAAT 480
 GCCATCCTTT CAACTTCTGA AGGTTTGCTT ACTGATAAAG AAGCACGCCA AAAGAATGTT 540
 GGTGGTGAGG TTATCGCTTA CGTTTGGTAA AATCAAGATA CAAAGCTCGT AAAGAACAAA 600
 GCAAAATTAG GAAGTTGGAG AAGTTTGT TT ACAAACAGGC CAACTTATCT ATTTTGCACA 660
 GTTCTTAGAG CGTGTTCAGT TCAGCTCTTG AGCTAAGTAA GTATCTGAAC CCCGTGAAAA 720
 CTGGCCGTGC TGGCATGTT GGGTAACAGG AGAATAAAA CATGTCACGT ATTGGTAATA 780
 AGTTCAGCTA AGGCCTTCGT AAAAGTT 807

1372

(2) INFORMATION FOR SEQ ID NO: 358:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 653 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 358:

```
CCCAGTATTT TTGTCCAAGC ACGACCAGAA AAGGATGATA CAGATCTGGA ATTGGCTCTC      60
TTAACCATCT tTGAACAAAA TCCTCAGGCT CAGGTCACTA TTTTCGGTGC CTGGGTGGC      120
CGTATTGACC ATATGTTGGC CAATGTCTTT CTGCCTAGCA ATCCTAAGTT GGCACCTAT      180
ATGCATCAAA TAGAAATTGA GGATGGGCAA AACTTGATTA CTTATTGTCC AGAAGGAATC      240
AGTCAGCTAG AACCTCGTTC AGACTACGAC TATCTAGCCT TTATGCCAGT TCGGGATAGC      300
CAAGTATGAG TTGACAGAGG AAAATTTTTT CTTTAAAAAA GTGTACGCTT CTAACGAATA      360
TATAGATAGG GAAGTGTCGG TAACTTGCCC AGATGGTTAT GTGGTCGTAC TGCATAGCAA      420
GGACAGGAGG TAGGATGGAA AGTTTACTTA TTCTATTATT AATTGCCAAT CTAGCTGGTC      480
TCTTTCTGAT TTGGCAAAGG CAGGATAGGC AGGAGAAACA CTTAAGTAAG AGCTTGGAGG      540
ATCAGGCAGA TCATTTGTCA GACCAGCTGG ATTACCGCTT TGACCAAGCC AGACAAGCCA      600
GCCAGTTAGA CCAAAAAGAT TTGGAAGTGG TTGTCAGCGA CCGTTTGCAA GAA          653
```

(2) INFORMATION FOR SEQ ID NO: 359:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 641 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 359:

```
CACCATGTGA TGTGACGCTG GCCACAGCTG TCAGAAATCT GGCGAGCCAT CGTGTGCAAT      60
GACTCTTCCC GATGTAATCT TGTTCATAGT CCTTTGATGA ATATGTTCAA GCTGTAGAAG      120
GTGCGCTTCC TGAACACTTA TCAACTGTTA CAGGCGAGTT GACCAGTCAG GAAACAGATG      180
GCTGGTACAC ACTTGCCAAC ACTTCTTCAT CCCGCATTTA CCTAAAACAA GCCTTCCAAG      240
AAAATAGCAA CCTCCTAGAG CAAGTGGTAG AACCTTGAC TATTATCACT GGTGGACACA      300
ACCACAAGGA CCAGTTGACC TATGCTTGA AAACACTTTT GCAGAATGCG CCACATGATA      360
GTATCTGTGG CTGTAGCCTG GACGAAGTTC ACCGCGAGAT GGAAACGCGT TTTGCCAAGG      420
```

1373

TCAACCAAGT AGGAAACTTT GTTAAAAGTA ACTTGCTCAA CGAGTGGAAG GGTAAAATTG 480
CTACGGATAA GGCTCAAAGT GACTATCTCT TTAAGTGCAT TAACACAGGC TTGCATGATA 540
AGGTCGATAC TGTCAGCACA GTGATTGATG TGGCGACTTG TGATTTCAG GAATTGCACC 600
CAACAGAAGG CTACAAAAG ATGGCTGCTC TTATCTTGCC G 641

(2) INFORMATION FOR SEQ ID NO: 360:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1958 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 360:

CCTCAAGGCC AATTGAAGG CTCTAAAACA ATGGAAAAGT GCTACACAGA TGTGACAGAA 60
TTTGCCATTC CAGCAAGTAC TCAAAAGCTT TACTTATCAC CAGTTTAGA TGGCTTTAAT 120
AGCGAAATTA TTGCTTTTAA TCTTTCGACT TCACCAACT TAGAACAAGT ACAAACAATG 180
TTAGAACAGG CATTCAAAGA GAAGCACTAC GAGAATACGA TTCTCCATAG TGACCAAGGC 240
TCCCAATATC AACACGATTC TTATCATCGG TTCCTAGAGA GTAAGGGAAT TCAAGCATCT 300
ATGTCACGCA AGGGCAACAG CCAAGACAAC GGTATGATGG AATCTTTCTT TGGCATTTTA 360
AAATCCGAAA TGTTTTATGG CTATGAGAAA ACATTTAAAT CACTTAACCA ATTGGAACAA 420
GCCATTATAG ACTATATTGA TTACTACAAC AACAAACGAA TTAAGGTAAA ACTAAAAGGA 480
CTTAGTCCTG TGCAGTACAG AACTAAATCC TTTGGATAAA TTAATTGTCT AACTTTTGG 540
GGTCAGTACA AAACCTCTGC TACTATGCGT TTTATTATTG AAAGACTTAT TGGACTTTCT 600
CTCAAATCGA GTTTTACTC AATTTTCTTA CTGATTGGG ATTGAAATTC CAATTAATTT 660
CTCTGAGTAG AGTGTCTTGA TATTGGCTTC ATCAACAGAG GCCTTATCAA TTTTACGTTT 720
CAAGAAAAAT TCTTGAATGG TTTCGATTTC AGGCTCACGA ATAGCACGGT GTTTGTTTGA 780
GATGAGGATT TCATAGTGAA GCGGAGCTTG GGTAAAAATA ACATCTGTAT TCCCTGCAGA 840
ATAAACCTCA ACAAGGGTTG CATCGGTACT TTCTAGCTGA CTTTTTACAA GTTGGCAGTG 900
TGAGTTTGTC GTATTGATAA GCTTCATAAT ATTTCTCCG ATTTTCTAAT TCTATTATAG 960
CACTTTTGA ATAAAGTCGC TTGATTATA CTCAATGAAA ATCAAAGAGC AAAC TAGGAA 1020
GCTAGCCGCA GGCTATACTT GAGTACGGTA AGGCGACGCT GACGTGGTTT GAATTTTATT 1080
TTCGAAGAGT ATTAGCCAAT CTTATGCTGT TTTTCCAAG ATTCAATGGC CCATTTATGG 1140

1374

CTACCACGTT TAAGGTTTTT GATAGCCTCG TCAATAGGGA ACCAGGCAAT ATGATTAAAG	1200
TTTTCTAGTG GCTTTTGTAC TTCTTTGAAA GGAGTTGCTT CATAGAGGTA GGCAGGATTG	1260
TAGTAGTAGG TATCACGATG ACGAGAATAG AAATATTCGT CAGCTTGTCG GTAATAGGTA	1320
CCAATTTCTG CTGTGAAACC AAGCTCTTCA ATCAACTCAT GCTTTAGGGC TTCCTGATGA	1380
TTTTCACCTG CTTCATTTT TCCACATGGT AGGAACCAAG CACCATTGCG TTCTTGAACA	1440
AGAACAATTT GTTTTGTTC AGGATTAGGG ATAAGTGCAT ATACGCCATA GCGAGCAATA	1500
TAGTCTGTAT TCACTTTTTT TCTCCGAAAG TTGGGTTTGC CATTGCATTT TCCTCATTAT	1560
CTAGTATCGT TATTATTATA GTGAAATGAA CCAAAAATAG TACACAATGT GGTATAATCT	1620
TCTTATGGCA TATTCAATAG ATTTTCGTAA AAAAGTTCTC TCTTATTGTG AGCGAACAGG	1680
TAGTATAACA GAAGCATCAC ACGTTTCCA AATCTCACGT AATACCATT ATGGCTGGTT	1740
AAAGCTAAAA GAGAAAACAG GAGAGCTAAA CCACCAAGTA AAAGGAATAA AACCAAGAAA	1800
GGTTGATAGA GATAGACTTA AAAACTATCT TACTGACAAT CCAGACGCTT ATTTGACTGA	1860
AATAGCTTCT GAATTTGGCT GTCATCCAAC TACCATCCAC TATGCGCTCA AAGCTATGGG	1920
TACACTCGAA AAAAAAAGA ACTACACCTA CTATGAAC	1958

(2) INFORMATION FOR SEQ ID NO: 361:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 851 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 361:

TATGAAATTA AGTTATGATG ATAAAGTTCA GATCTATGAA CTTAGAAAAC AAGGATATAG	60
CTTAGAGAAG CTTTCAAATA AATTTGGGAT AACAATTCT AATCTTAGGT ATATGATTAA	120
ATTGATTGAT CGTTACGGAA TAGAGTTCGT CAAAAAGGA AAAATCGTT ACTATTCTCC	180
TGATTTAAAA CAAGAAATGA TTAATAAAGT CTGACATGAA GGCTGGACTA AAGATAGAGT	240
TTCTCTTGAA TACGGTCTCC CAAGTCGTAC GATACTTCTT AACTGGCTAG CACAATACAG	300
GAAAAACGGG TATACTATTG TTGAGAAACC AAGAGGGAGA GTACCTGAGA GCGGAGAATG	360
CCATCCTAAA AAAGTTAAGA GAACTCCGAT TGAAGGAGGA AAAAGAGAAA GAAGAAAGAC	420
AGAAATTGTT TAAGAATTAA TGAAGTGGT TTCGTTAGAT CTTCTTTTAA AAGTCATTAA	480
ACTAGCTCGT TCGACCTACT ACTATCACTT GAAACAGCTA GATAAACCAG ATAAGGACCA	540
AGAGCTTAAA GCTGAAATTC AATCCATTTT TATCGAACAC AAAGGAAATT ATGCTTATCG	600

1375

TCGGATTAT TTAGAACTAA GAAATCCTGG TTATCTGGTA AATCATAAAA GAGTTCAAGG	660
CTTGATGAAA GTACTCAATt TACAAGCTAA AACGCGACAG AAACGAAAAT ATTCTTCTCA	720
TAAAGGAGAC GTTGGCAAGA AGGCAGAGAA TCTCATTCOA GGCCAATTG AAGGCTCTAA	780
AACAATGGAA CAGTGCTACA CAGATGTGAC AGAATTTGCC ATTCCAGTAA GTACTTAAAA	840
GCTTTACTTA T	851

(2) INFORMATION FOR SEQ ID NO: 362:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1168 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 362:

GGGTAGAATC GATATCTCCA ATGAGTTGGT tTAGCTGGTG AAAGTGTAAA AAGATTTTCGw	60
CCAATTCAAG GTTGAGGCAT CGCAACTAT GGACTGTTTC CTCGTCAGTT CTGGAAAGAA	120
AACGGGATAA GGTGGCTGT GAAGCAAGCT GCCCTCCTTC CAACAATTTT GGAAAGTAGG	180
CATCAGCTGA CAATTCTTTA CAAGCATAGT CCGTTCCATA ACCTGTTAAC AGTTGAAAGA	240
GGAAGTGGAC AAGGATATCT GAATCCGAAT AACGACAGTA GCGGCGTTGG TCATTCGTTA	300
CTAAATACTT AGAAATCCGC TCPTTTAGTT TCAACTGGGA AAAAAGTTCC TGAAAAAGA	360
TAAGACCACC ATACTGGGT AAATGACCTC CATCGAAAGA TAGTTGGTAA AAAGACTTGT	420
TTTGGAAGTG ATGATTTGGT AAAGTGTTC TGTGAGTTTC CTTCTTTTTT GTGTTTTTTT	480
CTACACTTAT ACCATAAAGG GGAACTCTT TTTGTCTAG TAAAAACAC CCATTGGGTG	540
AAAAAGAAA CCATCCAGGA TCTAAGCTAA GGCAAGGATT CTGGATGGTT TTTAGATTTG	600
GGGTGAATAA TTGGGGTTTT AGCTGCTTGC GGCAATCAG GTTCAGATAC AAAAAGTTAC	660
TCATCAACCT TTAGTGAAA TCCAAGTACA TTAACTATC TATTAGACTA TTACGCTGAT	720
AATATAGTCA ATTGAAACAA GAACAAGACA AAAGAGCCTC ATAAAAGGTA TTGCAACTTG	780
GTAATACCTT TTTGAGGTGC TTTTGTATAT GAGCCCATGT TTTCTCAATA GGATTGTACT	840
CAGGTGAGTA GGGAGGAAGA GGTAAAAGTT TATACCCAAA CTCTTCACAC AAGAGTTCTA	900
ACTTACCCAT TCTATGGAAT CTTGCATTAT CCATAATAAT AACCGATGGT GTGTTTAATG	960
TTGGTAAGAG AAATTTCTGA AACCAAGCTT CAAAAAGTC GTCGTCATC GTCTCTTCGT	1020
AAGTTATTGG AGCGATTAA TCACCATTTG TTAGACCTGC AACCAAAGAA ATCCTCTGAT	1080

1376
ATCTTCTTCC AGATACTTTG CCTCTTCTTA ACTGACCTTT TAATGAGCGA CCATATTCTC 1140
GATAAAAATA AGTATCGAAT CCTGTTTC 1168

(2) INFORMATION FOR SEQ ID NO: 363:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4483 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 363:

GTCAGCTTCA GCAAGCCCAT CAGCTTCTGA ATCTGCATCA ACCAGTGCCT CCGCTTCAGC 60
GTCAACCAGT GCGTCGGCTT CAGCGTCGAC AAGTGCTTCG GCTTCAGCAT CAACGAGTGC 120
GTCGGCCTCA GCAAGCGCAA GTACCTCAGC GTCAGCTTCC GCCTCAACCA GTGCGTCGGC 180
TTCAGCAAGC ACAAGTGCCT CAGCCTCAGC AAGTATCTCA GCGTCTGAAT CGGCATCAAC 240
GAGTGCGTCT GAGTCAGCAT CAACGAGTAC GTCAGCCTCA GCAAGCACAT CAGCTTCTGA 300
ATCTGCATCA ACCAGTGCCT CAGCCTCAGC ATCGACAAGC GCCTCAGCTT CAGCAAGTAC 360
CAGTGCTTCA GCCTCAGCGT CGACAAGTGC GTCGGCCTCA ACCAGTGCAT CTGAATCGGC 420
ATCAACCAGT GCGTCAGCCT CAGCAAGTAC TAGTGCATCA GCTTCAGCAT CAACGAGTGC 480
ATCGGCTTCA GCATCAACCA GTGCCTCGGC TTCAGCGTCA ACCAGTGCCT CAGCTTCAGC 540
AAGTACCAGT GCTTCAGTCT CAGCATCAAC AAGTGCTTCA GCCTCAGCAT CGACAAGTGC 600
CTCGGCTTCA GCAAGCACAT CAGCATCTGA ATCAGCGTCG ACAAGCGCCT CAGCTTCAGC 660
AAGTACCAGT GCGTCAGCCT CAGCGTCGAC AAGTGCGTCA GCCTCAGCAA GTACTAGTGC 720
ATCAGCTTCA GCATCAACGA GTGCATCGGC TTCGGCGTCA ACCAGTGCAT CAGAGTCAGC 780
AAGTACCAGT GCGTCAGCTT CCGCATCAAC AAGTGCTTCG GCTTCAGCAA GCACCAAGTGC 840
GTCGGCTTCA GCAAGTACTA GCGCCTCAGC CTCAGCCTCA ACCAGTGCCT CAGCCTCAGC 900
AAGTATCTCA GCGTCTGAAT CGGCATCAAC GAGTGCGTCC GCTTCAGCAA GTACTAGCGC 960
CTCAGCCTCA GCGTCAACAA GTGCATCGGC TTCAGCGTCA ACGAGTGCCT CTGAATCGGC 1020
ATCAACGAGT GCGTCCGCTT CAGCAAGTAC TAGCGCCTCA GCCTCAGCGT CAACAAGTGC 1080
ATCGGCTTCA GCATCAACGA GTGCGTCCGC TTCAGCAAGT ACTAGCGCCT CAGCCTCAGC 1140
GTCAACAAGT GCATCGGCTT CAGCGTCAAC GAGTGCGTCT GAGTCAGCAT CAACGAGTGC 1200
GTCAGCCTCA GCAAGCACAT CAGCTTCTGA ATCTGCATCA ACCAGTGCCT CAGCCTCAGC 1260
ATCGACAAGC GCCTCAGCTT CAGCAAGTAC CAGTGCGTCA GCTCAGCGTC GACAAGTGCs 1320

1377

TCrGCTTCAG CAAGTACCAG TGCCTCAGCC TCAGCAAGTA CCAGTGCKTC AGCCTCAGCG	1380
TCGACAAGTG CGTCGGCCTC AACCAGTGCA TCTGAATCGG CATCAACCAG TGCCTCAGCC	1440
TCAGCAAGTA CTAGCGCCTC AGCCTCAGCA TCAACGAGTG CGTCCGCTTC AGCAAGTACT	1500
AGTGCTCAG CTTCAGCAAG TACTAGCGCC TCAGCCTCAG CGTCGACAAG CGCCTCAGCT	1560
TCAGCAAGTA CCAGTGCGTC AGCCTCAGCG TCGACAAGTG CGTCGGCTTC AGCAAGTACC	1620
TCAGCGTCTG AATCAGCATC AACAAGTGCG TCGGCTTCAG CATCAACGAG TGCATCAGCT	1680
TCAGCATCAA CAAGTGCTTC AGCTTCAGCA AGTACCAGTG CGTCGGCTTC AGCATCAACG	1740
AGTGCTTCAG TCTCAGCGTC AACCAGTGCC TCTGAATCCG CATCAACAAG TGCTCGGCT	1800
TCAGCAAGCA CCAGTGCTTC GGCTTCAGCG TCAACGAGTG CGTCTGAGTC AGCATCAACG	1860
AGTGCGTCAC CTCAGCAAGC ACATCAGCTT CTGAATCTGC ATCAACCAGT GCGTCACTTC	1920
CGCATCAACA AGCGCCTCGG CCTCAGCAAG TACAAGTGCT TCAGCCTCAG CATCAACCAG	1980
TGCATCAGCT TCAGCCTCAA CAAGTGCTTC AGCCTCAGCG TCAACCAGTG CCTCGGCTTC	2040
AGCAAGTACC AGTGCGTCAG cTTCAGCAAG CACAAGTGCG TCAGCTTCAG CATCAACCAG	2100
TGCTTCGGCT TCGGCATCAA CAAGTGCTTC AGCATCAGCA TCAACGAGTG CGTCAcTCA	2160
GCAAGTACTA GTGCATCAGC ATCAGCATCA ACCAGTGCGT CAGCCTCAGC AAGTATCTCA	2220
GCGTCTGAAT CGGCATCAAC GAGTGATCA GCATCAGCAT CAACGAGTGC ATCGGCTTCA	2280
GCGTCAACCA GTGCATCAGT CTCAGCAAGC ACCAGTGCGT CGGCTTCAGC ATCAACCAGT	2340
GCCTCAGCCT CAGCAAGTAT CTCAGCGTCT GAATCGGCAT CAACGAGTGC GTCAGcTCA	2400
GCAAGTACTA GTGCATCAGC ATCAGCATCA ACGAGTGCGT CGGCTTCAGC AAGTACCAGC	2460
GCCTCAGCTT CAGCAAGCAC CAGTGCGTCA GCCTCAGCAA GTACCAGCGC CTCAGCCTCA	2520
GCAAGCACCA GTGCCTCAGC TTCAGCAAGT ACCAGTGCGT CAGCCTCAGC GTCGACAAGT	2580
GCGTCGGCTT CAGCAAGTAC CTCAGCGTCT GAATCAGCAT CAACGAGTGC ATCAGCTTCA	2640
GCATCAACAA GTGCTTCAGC TTCAGCAAGT ACCAGTGCGT CGGCTTCAGC ATCAACGAGT	2700
GCTTCAGTCT CAGCGTCAAC CAGTGCTCTT GAATCAGCAT CAACAAGTGC CTCGGCTTCA	2760
GCAAGCACCA GTGCGTCGGC TTCAGCAAGT ACTAGTGCGT CGGCTTCAGC ATCGACAAGT	2820
GCGTCTGAAT CGGCATCAAC GAGTGCTTCG GCTTCAGCAT CAACGAGTGC GTCAGCCTCA	2880
GCAAGCACAT CAGCTTCTGA ATCTGCATCA ACCAGTGCGT CCGCTTCAGC GTCAACCAGT	2940
GCGTCGGCTT CAGCGTCGAC AAGTGCTTCG GCTTCAGCAT CAACGAGTGC GTCGGCCTCA	3000
GCAAGCGCAA GTACCTCAGC GTCAGCTTCC GCCTCAACCA GTGCGTCCGC TTCAGCAAGC	3060

1378

ACAAGTGGCT CAGCCTCAGC AAGTATCTCA GCGTCTGAAT CGGCATCAAC GAGTGGCTCG	3120
GCCTCAGCAA GCGCAAGTAC CTCAGCGTCA GCTTCCGCCT CAACCAGTGC GTCGGCTTCA	3180
GCAAGCACAA GTGCGTCAGC CTCAGCAAGT ATCTCAGCGT CTGAATCGGC ATCAACGAGT	3240
GCGTCTGAGT CAGCATCAAC GAGTACGTCA GCCTCAGCAA GCACATCAGC TTCTGAATCG	3300
GCATCAACCA GTGCGTCAGC CTCAGCATCG ACAAGCGCCT CAGCTTCAGC AAGTACCAGT	3360
GCTTCAGCCT CAGCGTCGAC AAGTGGCTCG GCCTCAACCA GTGCATCTGA ATCGGCATCA	3420
ACCAGTGGCT CAGCCTCAGC AAGTACTAGT GCATCAGCTT CAGCATCAAC GAGTGCATCG	3480
GCTTCAGCAT CAACCAGTGC CTCGGCTTCA GCGTCAACCA GTGCGTCAGC TTCAGCAAGT	3540
ACCAGTGCTT CAGTCTCAGC ATCAACAAGT GCTTCAGCCT CAGCATCGAC AAGTGCCTCG	3600
GCTTCAGCAA GCACATCAGC ATCTGAATCA GCGTCGACAA GCGCCTCAGC TTCAGCAAGT	3660
ACCAGTGGCT CAGCCTCAGC GTCGACAAGT GCGTCAGCCT CAGCAAGTAC TAGTGCATCA	3720
GCTTCAGCAT CAACGAGTGC ATCGGCTTCG GCGTCAACCA GTGCATCAGA GTCAGCAAGT	3780
ACCAGTGGCT CAGCTTCCGC ATCAACAAGT GCCTCGGCTT CAGCAAGCAC CAGTGGCTCG	3840
GCTTCAGCAA GTACTAGCGC CTCAGCCTCA GCCTCAACCA GTGCGTCAGC CTCAGCAAGT	3900
ATCTCAGCGT CTGAATCGGC ATCAACGAGT GCGTCCGCTT CAGCAAGTAC TAGCGCCTCA	3960
GCCTCAGCGT CAACAAGTGC ATCGGCTTCA GCGTCAACGA GTGCGTCTGA ATCGGCATCA	4020
ACGAGTGGCT CCGCTTCAGC AAGTACTAGC GCCTCAGCCT CAGCGTCAAC AAGTGCATCG	4080
GCTTCAGCAT CAACGAGTGC GTCCGCTTCA GCAAGTACTA GCGCCTCAGC CTCAGCGTCA	4140
ACAAGTGCAT CCGGTTTCAGC GTCAACGAGT GCGTCTGAGT CAGCATCAAC GAGTGGCTCA	4200
CCTCAKCAAG CACATCAGCT TCTGAATCTG CATCAACCAG TGCCTCACTT CCGCATCAAC	4260
AAGCGCCTCG GCCTCAGCAA GTACAAGTGC TTCAGCCTCA GCATCAACCA GTGCATCAGC	4320
TTCAGCCTCA ACAAGTGCTT CAGCCTCAGC GTCAGACCAG TGCCTCGGCT TCAGCAAGTA	4380
CCAGTGGCTC ACTTCAGCAA GCACAAGTGC GTCAGCTTCA GCATCAACCA GTGCTTCGGC	4440
TTCCGGCATCA ACAAGTGCCT CAGCATCAGC ATCAACGAGT GCG	4483

(2) INFORMATION FOR SEQ ID NO: 364:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2550 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 364:

1379

GTACCTCAGC GTCCTTCCGC CTCAACCAGT GCGTCCGCTT CAGCAAGCAC AAGTGCGTCA	60
CCTCAGCAAG TATCTCAGCG TCTGAATCGG CATCAACGAG TCGTTCGGCC TCAGCAAGCG	120
CAAGTACCTC AGCGTCACTT CCGCCTCAAC CAGTGCCTCG GCTTCAGCAA GCACAAGTGC	180
GTCACTCAG CAAGTATCTC AGCGTCTGAA TCGGCATCAA CGAGTGCCTC TGAGTCAGCA	240
TCAACGAGTA CGTCAGCCTC AGCAAGCACA TCAGCTTCTG AATCGGCATC AACCAGTGCG	300
TCAGCCTCAG CATCGACAAG CGCCTCAGCT TCAGCAAGTA CCAGTGCTTC AGCCTCAGCG	360
TCGACAAGTG CGTCGGCCTC AACCAGTGCA TCTGAATCGG CATCAACCAG TCGTTCAGCC	420
TCAGCAAGTA CTAGTGATC AGCTTCAGCA TCAACGAGTG CATCGGCTTC AGCATCAACC	480
AGTGCCTCGG CTTTCAGCGTC AACCAGTGCG TCAGCTTCAG CAAGTACCAG TGCTTCAGTC	540
TCAGCATCAA CAAGTGCTTC AGCCTCAGCA TCGACAAGTG CCTCGGCTTC AGCAAGCACA	600
TCAGCATCTG AATCAGCGTC GACAAGTGG TCGGCCTCAA CCAGTGATC TGAATCGGCA	660
TCAACCAGTG CGTCAGCCTC AGCAAGTACT AGTGATCAG CTTTCAGATC AACGAGTGCA	720
TCGGCTTCGG CGTCAACCAG TGCATCAGAG TCAGCAAGTA CCAGTGCTTC AGCTTCCGCA	780
TCAACAAGTG CCTCGGCTTC AGCAAGCACA TCAGCATCTG AATCAGCGTC AACCAGTGCT	840
TCGGCTTCAG CAAGTACCAG TGCTTCAGCT TCAGCATCAA CCAGCGCCTC GGCCTCAGCA	900
AGCACCTCAG CTTTCTGAATC GGCCTCAACC AGCGCCTCGG CCTCAGCAAG CACCTCAGCT	960
TCTGAATCGG CCTCAACCAG CGCCTCAGCC TCAGCATCAA CGAGTGCTTC GGCTTCAGCA	1020
AGCACAAGCG CCTCGGCTTC AGCATCAACG AGTACGTCAG CTTTCAGCGTC AACCAGTGCT	1080
TCAGCCTCAG CATCAACAAG TCGTTCAGCC TCAGCAAGTA TCTCAGCGTC TGAATCGGCA	1140
TCAACGAGTG CGTCTGAGTC AGCATCAACG AGTACGTCAG CCTCAGCAAG CACAAGTGCT	1200
TCAGCCTCAG CAAGTATCTC AGCGTCTGAA TCGGCATCAA CGAGTGCTTC CGCTTCAGCA	1260
AGTACTAGCG CCTCAGCATC AGCGTCAACA AGTGCTTCGG CTTTCAGCGTC AACGAGTGCG	1320
TCTGAGTCAG CATCAACGAG TACGTCAGCC TCAGCAAGCA CATCAGCTTC TGAATCTGCA	1380
TCAACCAGTG CGTCAGCCTC AGCATCGACA AGCGCCTCAG CTTTCAGCAAG TACCAGTGCG	1440
TCAGCCTCAG CAAGTACCAG TGCTTCAGCC TCAGCGTCGA CAAGTGCTTC GGCCTCAACC	1500
AGTGATCTG AATCGGCATC AACCAGTGCG TCAGCTCAGC AAGTACTAGT GCATCAGCTT	1560
CAGCATCAAC GAGTGATCG GCTTCGGCGT CAACCAGTGC ATCAGAGTCA GCAAGTACCA	1620
GTGCGTCACT TCCGCATCAA CAAGTGCTTC GGCTTCAGCA AGCACATCAG CATCTGAATC	1680
AGCGTCAACC AGTGCTTCGG CTTTCAGCAAG TACCAGTGCT TCAGCTTCAG CATCAACCAG	1740

1380

CGCCTCGGCC TCAGCAAGCA CCTCAGCTTC TGAATCGGCC TCAACCAGCG CCTCGGCCTC	1800
AGCAAGCACC TCAGCTTCTG AATCGGCCTC AACCAGCGCC TCAGCCTCAG CATCAACGAG	1860
TGCTTCGGCT TCAGCAAGCA CAAGCGCCTC GGGTTCAGCA TCAACGAGTA CGTCAGCTTC	1920
AGCGTCAACC AGTGCTTCAG CCTCAGCATC AACAAGTGCG TCAGCCTCAG CAAGTATCTC	1980
AGCGTCTGAA TCGGCATCAA CGAGTGCGTC TGAGTCAGCA TCAACGAGTA CGTCAGCCTC	2040
AGCAAGCACC TCAGCTTCTG AATCGGCCTC AACCAGTGCG TCAGCCTCAG CATCGACAAG	2100
CGCCTCAGCT TCAGCAAGTA CCAGTGCTTC AGCCTCAGCG TCGACAAGTG CGTCGGCCTC	2160
AACCAGTGCA TCTGAATCGG CATCAACCAG TCGTCAGCC TCAGCAAGTA CTAGTGCATC	2220
GGCTTCAGCA TCAACCAGTG CCTCGGCCTC AGCGTCAACC AGTGCGTCAG CTTAGCAAG	2280
TACCAGTGCT TCAGTCTCAG CATCAACAAG TGCTTCAGCC TCAGCATCGA CAAGTGCCTC	2340
GGCTTCAGCA AGCACATCAG CATCTGAATC AGCGTCGACA AGCGCCTCAG CTTAGCAAG	2400
TACCAGTGCG TCAGCCTCAG CGTCGACAAG TCGTCAGCT ACAGCAAGTA CTAGTGCATC	2460
AGCTTCAGCA TCAACGAGTG CATCGGCTTC GCGCTCAACC AGTGCATCAG AGTCAGCAAG	2520
TACCAGTGCG TCAGTTCACG CATCAACAAG	2550

(2) INFORMATION FOR SEQ ID NO: 365:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1436 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 365:

ACCCAGCAAG TACTAGTGCA TCGGCTTCAG CAAGCACCAG TCGCTCGGCT TCAGCATCAA	60
CCAGTGCCCTC AGCCTCAGCA AGTATCTCAG CGTCTGAATC GGCATCAACG AGTGCGTCAC	120
CTCAGCAAGT ACTAGTGCAT CAGCATCAGC ATCAACGAGT GCATCGGCTT CAGCAAGTAC	180
CAGCGCCTCA GCTTCAGCAA GCACCAGTGC GTCAsCTCAG CAAGTACCAG CGCCTCAGCC	240
TCAGCAAGCA CCAGTGCCCTC AGCTTCAGCA AGTACCAGTG CGTCAGCCTC AGCGTCGACA	300
AGTGCGTCGG CTTAGCAAG TACCTCAGCG TCTGAATCAG CATCAACGAG TGCATCAGCT	360
TCAGCATCAA CAAGTGCTTC AGCTTCAGCA AGTATCTCAG CGTCTGAATC GGCATCAACG	420
AGTGCGTCCG CTTAGCAAG TACTAGCGCC TCAGCATCAG CGTCAACAAG TGCTTCGGCT	480
TCAGCGTCAA CGAGTGCGTC TGAGTCAGCA TCAACGAGTA CGTCAGCCTC AGCAAGCACA	540
TCAGCTTCTG AATCTGCATC AACCAGTGCG TCAGCCTCAG CATCGACAAG CGCCTCAGCT	600

1381

TCAGCAAGTA CCAGTGCCTC AGCCTCAGCA AGTACCAGTG CTTAGCCTC AGCGTCGACA	660
AGTGCCTCGG CCTCAACCAG TGCATCTGAA TCGGCATCAA CCAGTGCCTC AGCCTCAGCA	720
AGTACTAGCG CCTCAGCCTC AGCATCAACG AGTGCCTCGG CTTAGCAAG TACTAGTGCA	780
TCAGCTTCAG CAAGTACTAG CGCCTCAGCC TCAGCGTCGA CAAGCGCCTC AGCTTCAGCA	840
AGTACCAGTG CGTCAGCCTC AGCGTCGACA AGTGCCTCGG CTTAGCAAG TACCTCAGCG	900
TCTGAATCAG CATCAACAAG TGCCTCGGCT TCAGCATCAA CGAGTGCATC AGCTTCAGCA	960
TCAACAAGTG CTTAGCTTC AGCAAGTACC AGTGCCTCGG CTTAGCATC AACGAGTGCT	1020
TCAGTCTCAG CGTCAACCAG TGCCTCTGAA TCCGCATCAA CAAGTGCCTC GGCTTCAGCA	1080
AGCACCAGTG CTTGGCTTC AGCGTCAACG AGTGCCTCTG AGTCAGCATC AACGAGTGCG	1140
TCAGCCTCAG CAAGCACATC AGCTTCTGAA TCTGCATCAA CCAGTGCCTC AGCTTCCGCA	1200
TCAACAAGCG CCTCGGCCTC AGCAAGTACA AGTGCTTCAG CCTCAGCATC AACGAGTGCA	1260
TCAGCTTCAG CCTCAACAAG TGCTTCAGCC TCAGCGTCAA CCAGTGCCTC GGCTTCAGCA	1320
AGTACCAGTG CGTCAGCTTC AGCAAGCACA AGTGCCTCAG CTTAGCATC AACGAGTGCT	1380
TCGGCTTCGG CATCAACAAG TGCCTCAGCA TCAGCATCAA CGAGTGCCTC AGCCGG	1436

(2) INFORMATION FOR SEQ ID NO: 366:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 735 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 366:

GCAGTTGCCA CACCGTGCTG ACCAGCACCC GTTCCTGCGA TAATTTCTT TTTACCCATG	60
CGTWTGGCAA GCCAACTTG TCCTAAGGCA TTGTTAATCT TGTGGGCTCC TGTATGGTTA	120
AGGTCTTCCC GTTTGAGATA AATCTTGCTC CGCCAATATG CTGGGTCAAG TTTTTCGCT	180
AATAAGAGG AGTTTCAGT CCTACGTACT GCGCAAAAG CTGGTTAAT TCCTCTTGA	240
AACTTGGGTC TGCTGACTT TCACGGTAGG CTTCTCCAA CTCCAAAAC TGTGTATCA	300
ATGTTTCTGG GACAAAACG CCGCCGAATT TTCCGTAAAA TCCATCTTTA TTTGGTTCCT	360
GATATGCCAT GCTTTACCCT CTCTATAAAT CTTCTAATCT TTTCATGATC TTTTGTCCA	420
TCTGTCTCCA CTCCGCTCGA TACATCTACT GCATAGGGAG TAAAGTGTTG AATTGCTTTT	480
ACTACATTAT CTTCAATTAAG GCCACCTGCG ATAAAGAAGG GCTGTGCTAG TCCAGTCGTA	540

1382

TCCAGTTGAC CCCAATCAAA GGGCTGGCCA CTTCTGCGCA CAGGGGCATC AAAGAGTAGA	600
TAATCTGCCT GAGAATTGGG GACATGCCCC TTTCCATCTA CCTGCACAGC CTGAATACTG	660
GCACAAGGCA AATTCTCAAA TAAATCATCT GCCACCTGAC CGTGAAGTTG AACCAAGTCC	720
AAGCCGGGGA TCCTC	735

(2) INFORMATION FOR SEQ ID NO: 367:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1702 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 367:

TACTAGCGCC TCAGCCTCAG CGTCAACAAG TGCATCGGCT TCAGCATCAA CGAGTCCGTC	60
CGCTTCAGCA AGTACTAGCG CCTCAGCCTC AGCGTCAACA AGTGCATCGG CTTAGCGTC	120
AACGAGTGGC TCTGAGTCAG CATCAACGAG TGCCTCAGCC TCAGCAAGCA CATCAGCTTC	180
TGAATCTGCA TCAACCAAGT CGTCAGCCTC AGCATCGACA AGCGCCTCAG CTTAGCAAG	240
TACCAGTGGC TCAGCCTCAG CGTCGACAAG TGCCTCGGCT TCAGCAAGTA CCAGTCCGTC	300
AGCCTCAGCA AGTACCAGT CGTCAGCCTC AGCGTCGACA AGTGCCTCGG CCTCAACCAG	360
TGCATCTGAA TCGGCATCAA CCAGTCCGTC AGCCTCAGCA ACTACTACTC CTTAGCCTTC	420
AGCATCAACG AGTGCATCGG CTTAGCATC AACCAAGTGA TCAGAGTCAG CAAGTACCAG	480
TGCCTCAGCT TCCGCATCAA CAAGTGCCTC GGCTTCAGCA AGTACTAGCG CCTCAGCCTC	540
AGCGTCAACA AGTGCTTCAG CTTCCGCGTC AACCAGCGCC TCGGCCTCAG CAAGTATCTC	600
AGCGTCTGAA TCGGCATCAA CAAGTGCCTC GGCTTCAGCA TCAACGAGTG CATCAGTCTC	660
AGCAAGCACC AGTGCCTCGG CCTCAGCAAG CACCAGCGCG TCTGAATCCG CATCAACCAG	720
TGCCTCAGCT TCAGCAAGTA CCTCAGCATC TGAATCAGCA TCAACAAGTG CATCGGCTTC	780
AGCAAGCACA AGTGCTTCAG CCTCAGCAAG TATCTCAGCG TCTGAATCGG CATCAACGAG	840
TGCCTCCGCT TCAGCAAGTA CTAGCGCCTC AGCATCAGCG TCAACAAGTG CTTCCGCTTC	900
AGCGTCAACG AGTGCCTCTG AGTCAGCATC AACGAGTACG TCAGCCTCAG CAAGCACATC	960
AGCTTCTGAA TCTGCATCAA CCAGTCCGTC AGCCTCAGCA TCGACAAGCG CCTCAGCTTC	1020
AGCAAGTACC AGTGCCTCAG CCTCAGCAAG TACCAGTGCT TCAGCCTCAG CGTCGACAAG	1080
TGCCTCGGCC TCAACCAAGT CATCTGAATC GGCATCAACC AGTGCCTCAG CCTCAGCAAG	1140
TACTAGCGCC TCAGCCTCAG CATCAACGAG TGCCTCCGCT TCAGCAAGTA CTAGTGCATC	1200

1383

AGCATCAGCA TCAACGAGTG CATCGGCTTC AGCAAGTACC AGCGCCTCAG CTTAGCAAG	1260
CACCAGTGCG TCAGCCTCAG CAAGTACCAG CGCCTCAGCC TCAGCAAGCA CCAGTGCCTC	1320
AGCTTCAGCA AGTACCAGTG CGTCAGCCTC AGCGTCGACA AGTGCGTCGG CTTAGCAAG	1380
TACCTCAGCG TCTGAATCAG CATCAACGAG TGCATCAGCT TCAGCATCAA CAAGTGCTTC	1440
AGCTTCAGCA AGTACCAGTG CGTCGGCTTC AGCATCAACG AGTGCTTCAG TCTAGCGTC	1500
AACCAAGTCC TCTGAATCAG CATCAACAAG TGCCTCGGCT TCAGCAAGCA CCAGTGCCTC	1560
GGCTTCAGCA AGTACTAGTG CATCGGCTTC AGCATCGACA AGTGCGTCTG AATCGGCATC	1620
AACGAGTGCT TCGGCTTCAG CATCAACGAG TCGGTCAGCC TCAGCAAGCA CATCAGCTTC	1680
TGAATCTGCA TCAACCAAGTG CG	1702

(2) INFORMATION FOR SEQ ID NO: 368:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 941 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 368:

ACCAGTGCAT CAGCTTCAGC CTCAACAAGT GCTTCAGCCT CAGCGTCAAC CAGTGCCTCG	60
GCTTCAGCAA GTACCAGTGC GTCAGTTCAG CAAGCACAAG TCGCTCACTT CAGCATCAAC	120
CAGTGCTTCG GCTTCGGCAT CAACAAGTGC CTCAGCATCA GCATCAACGA GTGCGTCACC	180
TCAGCAAGTA CTAGTGCATC AGCATCAGCA TCAACCAAGTG CATCAGCCTC AGCAAGTATC	240
TCAGCGTCTG AATCGGCATC AACGAGTGCA TCAGCATCAG CATCAACGAG TGCATCGGCT	300
TCAGCGTCAA CCAGTGCATC AGTCTCAGCA AGCACCAGTG CGTCGGCTTC AGCATCAACG	360
AGTGCCTCAG CCTCAGCAAG TATCTCAGCG TCTGAATCGG CATCAACGAG TCGCTCAGCC	420
TCAGCAAGTA CTAGTGCATC GGCTTCAGCA AGCACCAGTG CGTCGGCTTC AGCATCAACC	480
AGTGCCTCAG CCTCAGCAAG TATCTCAGCG TCTGAATCGG CATCAACGAG TCGCTCAGCC	540
TCAGCAAGTA CTAGTGCATC AGCATCAGCA TCAACGAGTG CATCGGCTTC AGCAAGTACC	600
AGCGCCTCAG CTTAGCAAG CACCAGTGCG TCAGCCTCAG CAAGTACCAG CGCCTCAGCC	660
TCAGCAAGCA CCAGTGCCTC AGCTTCAGCA AGTACCAGTG CGTCAGCCTC AGCGTCGACA	720
AGTGCGTCGG CTTAGCAAG TACCTCAGCG TCTGAATCAG CATCAACGAG TGCATCAGCT	780
TCAGCATCAA CAAGTGCTTC AGCTTCAGCA AGTACCAGTG CGTCGGCTTC AGCATCAACG	840

1370

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 355:

CCGACAAGCA ATATTAAAA GAGTAACTA TTAAC TAGTT AATTAACCGG TTTATTACTT	60
TATAGTGAAT CAAATATACT TAAGAAAAGA GGAAAGAATG AAAATTAATA AAAATATCT	120
AGCAGGTTCA GTGGCAGTCC TTGCCCTAAG TGTTTGTTC TATGAGCTTG GACGTTACCA	180
AGCTGGTCAG GATAAGAAAG AGTCTAATCG AGTTGCTTAT ATAGATGGTG ATCAGGCTGG	240
TCAAAAGGCA GAAAACTGA CACCAGATGA AGTCAGTAAG AGGGAGGGGA TCAACGCCGA	300
ACAAATTGTT ATCAAGATTA CGGATCAAGG TTATGTGACC TCTCATGGAG ACCATTATCA	360
TTACTATAAT GGCAAGGTTT CTTATGATGC CATCATCAGT GAAGAGCTCC TCATGAAAGA	420
TCCGAATTAT CAGTTGAAGG ATTCAGACAT TGTCAATGAA ATCAAGGGTG GTTATGTCAT	480
TAAGGTAAAC GGTAATACT ATGTTTACCT TAAGGATGCA GCTCATGCGG ATAATATTCG	540
GACAAAAGAA GAGATTAAAC GTCAGAAGCA GGAACGCAGT CATAATCATA ACTCAAGAGC	600
AGATAATGCT GTTGCTGCAG CCAGAGCCCA AGGACGTTAT ACAACGGATG ATGGGTATAT	660
CTTCAATGCA TCTGATATCA TTGAGGACAC GGGTGATGCT TATATCGTTC CTCACGGCGA	720
CCATTACCAT TACATTCCTA AGAATGAGTT ATCAGCTAGC GAGTTAGCTG CTGCAGAAGC	780
CTATTGGAAT GGGAAGCAGG GATCTCGTCC TTCTTCAAGT TCTAGTTATA ATGCAAATCC	840
AGCTCAACCA AGATTGTCAG AGAACCACAA TCTGACTGTC ACTCCAACCT ATCATCAAAA	900
TCAAGGGGGA AACATTTCAA GCCTTTTACG TGAATTGTAT GCTAACCCCTT ATCAGAACGC	960
CATGTGGGAT CTG	973

(2) INFORMATION FOR SEQ ID NO: 356:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 843 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 356:

GGTCGCATCT GCAATATCTG TCGCCTCCAC ATAAGCGACA CCAGCCTTGT CTGCTGCCCG	60
TTTGACACGT TCTGCAGATT GACCCAGGAT GACCATCTTC TTGAGTCCAG TAATGTCTGG	120
CACCAATTCTG TCAAACCTCAT TGCCACGGTC CAAACCACCT GCAATCAAGA CGACCTTGCT	180
GTGTGCAAAAT CCTGACAAGC TTTTGTGAGTA GCCAAGATAT TAGTTGATTT ACTGTCGTTA	240
TAGAATTAA CACsCTTGAT GTCATCCACA AACTGGAGAC GGTGTTTGAC ACCACCGAAG	300
GCTGAAAGAG TTTCCTTGAT GGTGTTGATTG TCCACATCAC GAAGCTTGGC TACAGCAATA	360

WO 98/18931

1371

GTCGCAAGGG CATTTTCCAC ATTGTGGCTA CCTGGAACAC CGATTTTCATT CGCTGCCATG 420
ACTACTTCAC CACGGAAGTA GAGTTGACCA TCTTCCAGAT AAGCTCCATC AACCTTTTCA 480
AGTGTTGAAA ATGGTACAAC AGTGGCTTCT GTCTTGGAAG TCAAGTCTTT TGCCAAGTCT 540
TGATTAAAGT TCAAGACAAG GAAATCAGCT GCTGTCATCT TGTTCCTGGAT ATTCCACTTG 600
GCTGCTACAT ATTCCGAAAA TGACCCATGG TAGTCGATAT GAGTTGGCAT GAGGTTGGTA 660
ATAACCGCAA TCTCTGGATG GAATTCCTGA ACACCCATGA GTTGAAAGA AGAAAGTTCC 720
ATAACAAGCG TGTCCTTATC TGATGCTATT TGAGCAACCT GACTAGCTGG ATAGCCGATA 780
TTCCCTGATA AAAGACCATG TTGGCCAGCA GCAGTCAAAA CTCCCGGGn TCCTCTAGAG 840
TCG 843

(2) INFORMATION FOR SEQ ID NO: 357:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 807 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 357:

TTTTTTTAT ATTTTITTA TTTATTATTT TTTGGCAAAA AAGACCAATT TGCTTTGGAG 60
CATTGCTTCT GCATTAAATT GTCTATTTT GTCGTGCTG TTACGCTCTT TGTATCATGT 120
ATTAAGTAGC AAGTGCAACT TGCAAACTAC TAGTAAGAGG AGAAAAACAA AATGGTTATG 180
ACTGACCCAA TCGCAGACTT CCTAACTCGT ATTCGTAATG CTAACCAAGC TAAACACGAA 240
GTACTTGAAG TACCTGCATC AAACATCAAA AAAGGGATTG CTGAAATCCT TAAACGCGAA 300
GGTTTTGTAA AAAACGTTGA AATCATTGAA GATGACAAAC AAGGCGTCAT CCGTGATTTT 360
CTTAAATACG GACCAAAATGG TGAGAAAGTT ATCACTAACT TGAACGTGT TTCTAAACCA 420
GGACTTCGTG TCTACAAAAA ACGTGAAGAC CTTCCAAAAG TTCTTAACGG ACTTGGAATT 480
GCCATCCTTT CAACTTCTGA AGGTTTGCTT ACTGATAAAG AAGCACGCCA AAAGAATGTT 540
GGTGGTGAGG TTATCGCTTA CGTTTGGTAA AATCAAGATA CAAAGCTCGT AAAGAACAAA 600
GCAAAATTAG GAAGTTGGAG AAGTTTGTGTT ACAACAGGC CAACTTATCT ATTTTGCACA 660
GTTCTTAGAG CGTGTTCAGT TCAGCTCTTG AGCTAAGTAA GTATCTGAAC CCCGTGAAAA 720
CTGGCCGTGC TGGCATGTTT GGGTAACAGG AGAATAATAA CATGTCACGT ATTGGTAATA 780
AGTTCAGCTA AGGCCTTCGT AAAAGTT 807

1372

(2) INFORMATION FOR SEQ ID NO: 358:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 653 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 358:

CCCAGTATTT TTGTCCAAGC ACGACCAGAA AAGGATGATA CAGATCTGGA ATTGGCTCTC	60
TTAACCATCT tTGAACAAAA TCCTCAGGCT CAGGTCATA TTTTCGGTGC CTGGGTGGC	120
CGTATTGACC ATATGTTGGC CAATGTCTTT CTGCCTAGCA ATCCTAAGTT GGCACCTAT	180
ATGCATCAAA TAGAAATTGA GGATGGGCAA AACTTGATTA CTTATTGTCC AGAAGGAATC	240
AGTCAGCTAG AACCTCGTTC AGACTACGAC TATCTAGCCT TTATGCCAGT TCGGGATAGC	300
CAAGTATGAG TTGACAGAGG AAAATTTTTT CTTAAAAAA GTGTACGCTT CTAACGAATA	360
TATAGATAGG GAAGTGTCGG TAACTGCCC AGATGGTTAT GTGGTCGTAC TGCATAGCAA	420
GGACAGGAGG TAGGATGGAA AGTTTACTTA TTCTATTATT AATTGCCAAT CTAGCTGGTC	480
TCTTCTGAT TTGGCAAAGG CAGGATAGGC AGGAGAAACA CTTAAGTAAG AGCTTGGAGG	540
ATCAGGCAGA TCATTTGTCA GACCAGCTGG ATTACCGCTT TGACCAAGCC AGACAAGCCA	600
GCCAGTTAGA CCAAAAAGAT TTGGAAGTGG TTGTCAGCGA CCGTTTGCAA GAA	653

(2) INFORMATION FOR SEQ ID NO: 359:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 641 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 359:

CACCATGTGA TGTGACGCTG GCCACAGCTG TCAGAAATCT GGCGAGCCAT CGTGTGCAAT	60
GACTCTTCCC GATGTAATCT TGTTCATAGT CCTTGATGA ATATGTTCAA GCTGTAGAAG	120
GTGCGCTTCC TGAACACTTA TCAACTGTTA CAGGCGAGTT GACCAGTCAG GAAACAGATG	180
GCTGGTACAC ACTTGCCAAC ACTTCTTCAT CCCGCATTTA CCTAAAACAA GCCTTCCAAG	240
AAAATAGCAA CCTCCTAGAG CAAGTGGTAG AACCTTGAC TATTATCACT GGTGGACACA	300
ACCACAAGGA CCAGTTGACC TATGCTTGA AAACACTTTT GCAGAATGCG CCACATGATA	360
GTATCTGTGG CTGTAGCGTG GACGAAGTTC ACCGCGAGAT GGAAACGCGT TTTGCCAAGG	420

1373

TCAACCAAGT AGGAAACTTT GTTAAAAGTA ACTTGCTCAA CGAGTGGAAG GGTAAAATTG 480
CTACGGATAA GGCTCAAAGT GACTATCTCT TTAAGTGCAT TAACACAGGC TTGCATGATA 540
AGGTCGATAC TGTCAGCACA GTGATTGATG TGGCGACTTG TGATTTCAG GAATTGCACC 600
CAACAGAAGG CTACAAAAG ATGGCTGCTC TTATCTTGCC G 641

(2) INFORMATION FOR SEQ ID NO: 360:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1958 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 360:

CCTCAAGGCC AATTGAAGG CTCTAAAACA ATGGAAAAGT GCTACACAGA TGTGACAGAA 60
TTTGCCATTC CAGCAAGTAC TCAAAAGCTT TACTTATCAC CAGTTTTAGA TGGCTTTAAT 120
AGCGAAATTA TTGCTTTTAA TCTTTCGACT TCACCAACT TAGAACAAGT ACAAACAATG 180
TTAGAACAGG CATTCAAAGA GAAGCACTAC GAGAATACGA TTCTCCATAG TGACCAAGGC 240
TGGCAATATC AACACGATTC TTATCATCGG TTCCTAGAGA GTAAGGGAAT TCAAGCATCT 300
ATGTCACGCA AGGGCAACAG CCAAGACAAC GGTATGATGG AATCTTTCCT TGGCATTTTA 360
AAATCCGAAA TGTTTTATGG CTATGAGAAA ACATTTAAAT CACTTAACCA ATTGGAACAA 420
GCCATTATAG ACTATATTGA TTAACAAC AACAACGAA TTAAGGTAAA ACTAAAAGGA 480
CTTAGTCCTG TGCAGTACAG AACTAAATCC TTTGGATAAA TTAATTGTCT AACTTTTGG 540
GGTCAGTACA AAACCTCTGC TACTATGCGT TTTATTATTG AAAGACTTAT TGGACTTTCT 600
CTCAAATCGA GTTTTACTC AATTTTCTTA CTGATTGGG ATTGAAATTC CAATTAATTT 660
CTCTGAGTAG AGTGTCTTGA TATTGGCTTC ATCAACAGAG GCCTTATCAA TTTTACGTTT 720
CAAGAAAAAT TCTTGAATGG TTTCGATTTC AGGCTCACGA ATAGCACGGT GTTTGTTTGA 780
GATGAGGATT TCATAGTGAA GCGGAGCTTG GGTAAAAATA ACATCTGTAT TCCCTGCAGA 840
ATAAACCTCA ACAAGGGTTG CATCGGTACT TTCTAGCTGA CTTTTTACAA GTTGCAGTG 900
TGAGTTTGTC GTATTGATAA GCTTCATAAT ATTTCCTCCG ATTTTCTAAT TCTATTATAG 960
CACTTTTGA ATAAAGTCGC TTGATTATA CTCAATGAAA ATCAAAGAGC AACTAGGAA 1020
GCTAGCCGCA GGTATACTT GAGTACGTA AGGCGACGCT GACGTGGTTT GAATTTTATT 1080
TTCGAAGAGT ATTAGCCAAT CTTATGCTGT TTTTCCAAG ATTCAATGGC CCATTTATGG 1140

1374

CTACCACGTT TAAGTTTTT GATAGCCTCG TCAATAGGGA ACCAGGCAAT ATGATTAAAG	1200
TTTTCTAGTG GCTTTTGTAC TTCTTTGAAA GGAGTTGCTT CATAGAGGTA GGCAGGATTG	1260
TAGTAGTAGG TATCACGATG ACGAGAATAG AAATATTCGT CAGCTTGTCG GTAATAGGTA	1320
CCAATTTCTG CTGTGAAACC AAGCTCTTCA ATCAACTCAT GCTTTAGGGC TTCCTGATGA	1380
TTTTCACCTG CTTCATTTT TCCACATGGT AGGAACCAAG CACCATTGGG TTCTTGAACA	1440
AGAACAATTT GTTTTGTTC AGGATTAGGG ATAAGTGCAT ATACGCCATA GCGAGCAATA	1500
TAGTCTGTAT TCACTTTTTT TCTCCGAAAG TTGGGTTTGC CATTGCATTT TCCTCATTAT	1560
CTAGTATCGT TATTATTATA GTGAAATGAA CCAAAAATAG TACACAATGT GGTATAATCT	1620
TCTTATGGCA TATTCAATAG ATTTTCGTAA AAAAGTTCTC TCTTATTGTG AGCGAACAGG	1680
TAGTATAACA GAAGCATCAC ACGTTTCCA AATCTCACGT AATACCATTG ATGGCTGGTT	1740
AAAGCTAAAA GAGAAAACAG GAGAGCTAAA CCACCAAGTA AAAGGAATAA AACCAAGAAA	1800
GGTTGATAGA GATAGACTTA AAAACTATCT TACTGACAAT CCAGACGCTT ATTTGACTGA	1860
AATAGCTTCT GAATTTGGCT GTCATCCAAC TACCATCCAC TATGCGCTCA AAGCTATGGG	1920
TACACTCGAA AAAAAAAGA ACTACACCTA CTATGAAC	1958

(2) INFORMATION FOR SEQ ID NO: 361:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 851 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 361:

TATGAAATTA AGTTATGATG ATAAAGTTCA GATCTATGAA CTTAGAAAAC AAGGATATAG	60
CTTAGAGAAG CTTTCAAATA AATTTGGGAT AACAATTCT AATCTTAGGT ATATGATTAA	120
ATTGATTGAT CGTTACGGAA TAGAGTTCGT CAAAAAAGGA AAAAATCGTT ACTATTCTCC	180
TGATTTAAAA CAAGAAATGA TTAATAAAGT CTGACATGAA GGCTGGACTA AAGATAGAGT	240
TTCTCTTGAA TACGGTCTCC CAAGTCGTAC GATACTTCTT AACTGGCTAG CACAATACAG	300
GAAAAACGGG TATACTATTG TTGAGAAACC AAGAGGGAGA GTACCTGAGA GCGGAGAATG	360
CCATCCTAAA AAAGTTAAGA GAACTCCGAT TGAAGGAGGA AAAAGAGAAA GAAGAAAGAC	420
AGAAATTGTT TAAGAATTAA TGAAGTGGTT TTCGTTAGAT CTTCTTTTAA AAGTCATTAA	480
ACTAGCTCGT TCGACCTACT ACTATCACTT GAAACAGCTA GATAAACCAAG ATAAGGACCA	540
AGAGCTTAAA GCTGAAATTC AATCCATTTT TATCGAACAC AAAGGAAATT ATGCTTATCG	600

1375

TCGGATTAT TTAGAACTAA GAAATCGTGG TTATCTGGTA AATCATAAAA GAGTTCAAGG	660
CTTGATGAAA GTACTCAATt TACAAGCTAA AACGCGACAG AAACGAAAAT ATTCTTCTCA	720
TAAAGGAGAC GTTGGCAAGA AGGCAGAGAA TCTCATTCOA GGCCAATTTG AAGGCTCTAA	780
AACAATGGAA CAGTGCTACA CAGATGTGAC AGAATTTGCC ATTCCAGTAA GTACTTAAAA	840
GCTTTACTTA T	851

(2) INFORMATION FOR SEQ ID NO: 362:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1168 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 362:

GGGTAGAATC GATATCTCCA ATGAGTTGGT tTAGCTGGTG AAAGTGTAAA AAGATTTTCGw	60
CCAATTCAG GTTGAGGCAT CGAAACTAT GGAAGTTTC CTCGTCAGTT CTGGAAGAA	120
AACGGGATAA GGTGGCTGT GAAGCAAGCT GCCCTCCTTC CAACAATTTT GGAAAGTAGG	180
CATCAGCTGA CAATTCTTTA CAAGCATAGT CCGTTCCATA ACCTGTTAAC AGTTGAAAGA	240
GGAAGTGGAC AAGGATATCT GAATCCGAAT AACGACAGTA GCGGCGTTGG TCATTCGTTA	300
CTAAATACTT AGAAATCCGC TCTTTTAGTT TCAACTGGGA AAAAAGTTCC TGAAAAAGA	360
TAAGACCACC ATACTGGGTT AAATGACCTC CATCGAAAGA TAGTTGGTAA AAAGACTTGT	420
TTTGGAAGTG ATGATTTGGT AAAGTGTCA TGTGAGTTTC CTTTCTTTT GTGTTTTTT	480
CTACACTTAT ACCATAAAGG GGAAACTCTT TTTTGTCTAG TAAAAACAC CCATTGGGTG	540
AAAAAAGAAA CCATCCAGGA TCTAAGCTAA GGCAAGGATT CTGGATGGTT TTTAGATTG	600
GGGTGAATAA TTGGGGTTTT AGCTGCTTGC GGCAATCAG GTTCAGATAC AAAAAGTTAC	660
TCATCAACCT TTAGTGAAA TCCAAGTACA TTAACTATC TATTAGACTA TTACGCTGAT	720
AATATAGTCA ATTGAACAA GAACAAGACA AAAGAGCCTC ATAAAAGGTA TTGCAACTG	780
GTAATACCTT TTTGAGGTGC TTTTGTATAT GAGCCCATGT TTTCTCAATA GGATTGTACT	840
CAGGTGAGTA GGGAGGAAGA GGTAAAAGTT TATACCCAAA CTCTTCACAC AAGAGTTCTA	900
ACTTACCCAT TCTATGGAAT CTTGCATTAT CCATAATAAT AACCGATGGT GTGTTTAATG	960
TTGGTAAGAG AAATTTCTGA AACCAAGCTT CAAAAAGTC GCTCGTCATC GTCTCTTCGT	1020
AAGTTATGG AGCGATTAAC TCACCATTTG TTAGACCTGC AACCAAGAA ATCCTCTGAT	1080

1376
ATCTTCTTCC AGATACTTTG CCTCTTCTTA ACTGACCTTT TAATGAGCGA CCATATTCTC 1140
GATAAAAATA AGTATCGAAT CCTGTTTC 1168

(2) INFORMATION FOR SEQ ID NO: 363:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4483 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 363:

GTCAGCTTCA GCAAGCCCAT CAGCTTCTGA ATCTGCATCA ACCAGTGCCT CCGCTTCAGC 60
GTCAACCACT GCGTCGGCTT CAGCGTCGAC AAGTGCTTCG GCTTCAGCAT CAACGAGTGC 120
GTGGGCTTCA GCAAGCGCAA GTACCTCAGC GTCAGCTTCC GCCTCAACCA GTGCGTCGGC 180
TTCAGCAAGC ACAAGTGCGT CAGCCTCAGC AAGTATCTCA GCGTCTGAAT CGGCATCAAC 240
GAGTGCGTCT GAGTCAGCAT CAACGAGTAC GTCAGCCTCA GCAAGCACAT CAGCTTCTGA 300
ATCTGCATCA ACCAGTGCCT CAGCCTCAGC ATCGACAAGC GCCTCAGCTT CAGCAAGTAC 360
CAGTGCTTCA GCCTCAGCGT CGACAAGTGC GTCGGCTTCA ACCAGTGCAT CTGAATCGGC 420
ATCAACCACT GCGTCAGCCT CAGCAAGTAC TAGTGCATCA GCTTCAGCAT CAACGAGTGC 480
ATCGGCTTCA GCATCAACCA GTGCCTCGGC TTCAGCGTCA ACCAGTGCCT CAGCTTCAGC 540
AAGTACCAGT GCTTCAGTCT CAGCATCAAC AAGTGCTTCA GCCTCAGCAT CGACAAGTGC 600
CTCGGCTTCA GCAAGCACAT CAGCATCTGA ATCAGCGTCG ACAAGCGCCT CAGCTTCAGC 660
AAGTACCAGT GCGTCAGCCT CAGCGTCGAC AAGTGCGTCA GCCTCAGCAA GTACTAGTGC 720
ATCAGCTTCA GCATCAACGA GTGCATCGGC TTCGGCGTCA ACCAGTGCAT CAGAGTCAGC 780
AAGTACCAGT GCGTCAGCTT CCGCATCAAC AAGTGCTTCG GCTTCAGCAA GCACCAAGTGC 840
GTCGGCTTCA GCAAGTACTA GCGCCTCAGC CTCAGCCTCA ACCAGTGCCT CAGCCTCAGC 900
AAGTATCTCA GCGTCTGAAT CGGCATCAAC GAGTGCGTCC GCTTCAGCAA GTACTAGCGC 960
CTCAGCCTCA GCGTCAACAA GTGCATCGGC TTCAGCGTCA ACGAGTGCCT CTGAATCGGC 1020
ATCAACGAGT GCGTCCGCTT CAGCAAGTAC TAGCGCCTCA GCCTCAGCGT CAACAAGTGC 1080
ATCGGCTTCA GCATCAACGA GTGCGTCCGC TTCAGCAAGT ACTAGCGCCT CAGCCTCAGC 1140
GTCAACAAGT GCATCGGCTT CAGCGTCAAC GAGTGCGTCT GAGTCAGCAT CAACGAGTGC 1200
GTCAGCCTCA GCAAGCACAT CAGCTTCTGA ATCTGCATCA ACCAGTGCCT CAGCCTCAGC 1260
ATCGACAAGC GCCTCAGCTT CAGCAAGTAC CAGTGCGTCA GCTCAGCGTC GACAAGTGCs 1320

1377

TCrGCTTCAG CAAGTACCAG TGCCTCAGCC TCAGCAAGTA CCAGTGCKTC AGCCTCAGCG	1380
TCGACAAGTG CGTCGGCCTC AACCAGTGCA TCTGAATCGG CATCAACCAG TGCCTCAGCC	1440
TCAGCAAGTA CTAGCGCCTC AGCCTCAGCA TCAACGAGTG CGTCCGCTTC AGCAAGTACT	1500
AGTGCAATCAG CTTCAAGCAAG TACTAGCGCC TCAGCCTCAG CGTCGACAAG CGCCTCAGCT	1560
TCAGCAAGTA CCAGTGCGTC AGCCTCAGCG TCGACAAGTG CGTCGGCTTC AGCAAGTACC	1620
TCAGCGTCTG AATCAGCATC AACAAGTGCG TCGGCTTCAG CATCAACGAG TGCATCAGCT	1680
TCAGCATCAA CAAGTGCTTC AGCTTCAGCA AGTACCAGTG CGTCGGCTTC AGCATCAACG	1740
AGTGCTTCAG TCTCAGCGTC AACCAGTGCC TCTGAATCCG CATCAACAAG TGCCTCGGCT	1800
TCAGCAAGCA CCAGTGCTTC GGCTTCAGCG TCAACGAGTG CGTCTGACTC AGCATCAACG	1860
AGTGCGTCAC CTCAGCAAGC ACATCAGCTT CTGAATCTGC ATCAACCAGT GCGTCACTTC	1920
CGCATCAACA AGCGCCTCGG CCTCAGCAAG TACAAGTGCT TCAGCCTCAG CATCAACCAG	1980
TGCATCAGCT TCAGCCTCAA CAAGTGCTTC AGCCTCAGCG TCAACCAGTG CCTCGGCTTC	2040
AGCAAGTACC AGTGCGTCAG cTTCAAGCAAG CACAAGTGCG TCAGCTTCAG CATCAACCAG	2100
TGCTTCGGCT TCGGCATCAA CAAGTGCCCTC AGCATCAGCA TCAACGAGTG CGTCAcCTCA	2160
GCAAGTACTA GTGCATCAGC ATCAGCATCA ACCAGTGCGT CAGCCTCAGC AAGTATCTCA	2220
GCGTCTGAAT CGGCATCAAC GAGTGATCA GCATCAGCAT CAACGAGTGC ATCGGCTTCA	2280
GCGTCAACCA GTGCATCAGT CTCAGCAAGC ACCAGTGCGT CGGCTTCAGC ATCAACCAGT	2340
GCCTCAGCCT CAGCAAGTAT CTCAGCGTCT GAATCGGCAT CAACGAGTGC GTCAGcCTCA	2400
GCAAGTACTA GTGCATCAGC ATCAGCATCA ACGAGTGCGT CGGCTTCAGC AAGTACCAGC	2460
GCCTCAGCTT CAGCAAGCAC CAGTGCGTCA GCCTCAGCAA GTACCAGCGC CTCAGCCTCA	2520
GCAAGCACCA GTGCCTCAGC TTCAGCAAGT ACCAGTGCGT CAGCCTCAGC GTCGACAAGT	2580
GCGTCGGCTT CAGCAAGTAC CTCAGCGTCT GAATCAGCAT CAACGAGTGC ATCAGCTTCA	2640
GCATCAACAA GTGCTTCAGC TTCAGCAAGT ACCAGTGCGT CGGCTTCAGC ATCAACGAGT	2700
GCTTCAGTCT CAGCGTCAAC CAGTGCCTCT GAATCAGCAT CAACAAGTGC CTCGGCTTCA	2760
GCAAGCACCA GTGCGTCGGC TTCAGCAAGT ACTAGTGCGT CGGCTTCAGC ATCGACAAGT	2820
GCGTCTGAAT CGGCATCAAC GAGTGCTTCG GCTTCAGCAT CAACGAGTGC GTCAGCCTCA	2880
GCAAGCACAT CAGCTTCTGA ATCTGCATCA ACCAGTGCGT CCGCTTCAGC GTCAACCAGT	2940
GCGTCGGCTT CAGCGTCGAC AAGTGCTTCG GCTTCAGCAT CAACGAGTGC GTCGGCCTCA	3000
GCAAGCGCAA GTACCTCAGC GTCAGCTTCC GCCTCAACCA GTGCGTCCGC TTCAGCAAGC	3060

1378

ACAAGTGCGT CAGCCTCAGC AAGTATCTCA GCGTCTGAAT CGGCATCAAC GAGTGCGTCG	3120
GCCTCAGCAA GCGCAAGTAC CTCAGCGTCA GCTTCCGCCT CAACCAGTGC GTCGGCTTCA	3180
GCAAGCACAA GTGCGTCAGC CTCAGCAAGT ATCTCAGCGT CTGAATCGGC ATCAACGAGT	3240
GCGTCTGAGT CAGCATCAAC GAGTACGTCA GCCTCAGCAA GCACATCAGC TTCTGAATCG	3300
GCATCAACCA GTGCGTCAGC CTCAGCATCG ACAAGCGCCT CAGCTTCAGC AAGTACCAGT	3360
GCTTCAGCCT CAGCGTCGAC AAGTGCGTCG GCCTCAACCA GTGCATCTGA ATCGGCATCA	3420
ACCAAGTGCGT CAGCCTCAGC AAGTACTAGT GCATCAGCTT CAGCATCAAC GAGTGCATCG	3480
GCTTCAGCAT CAACCAGTGC CTCGGCTTCA GCGTCAACCA GTGCGTCAGC TTCAGCAAGT	3540
ACCAAGTGCTT CAGTCTCAGC ATCAACAAGT GCTTCAGCCT CAGCATCGAC AAGTGCCTCG	3600
GCTTCAGCAA GCACATCAGC ATCTGAATCA GCGTCGACAA GCGCCTCAGC TTCAGCAAGT	3660
ACCAAGTGCGT CAGCCTCAGC GTCGACAAGT GCGTCAGCCT CAGCAAGTAC TAGTGCATCA	3720
GCTTCAGCAT CAACGAGTGC ATCGGCTTCG GCGTCAACCA GTGCATCAGA GTGAGCAAGT	3780
ACCAAGTGCGT CAGCTTCCGC ATCAACAAGT GCCTCGGCTT CAGCAAGCAC CAGTGCGTCG	3840
GCTTCAGCAA GTACTAGCGC CTCAGCCTCA GCCTCAACCA GTGCGTCAGC CTCAGCAAGT	3900
ATCTCAGCGT CTGAATCGGC ATCAACGAGT GCGTCCGCTT CAGCAAGTAC TAGCGCCTCA	3960
GCCTCAGCGT CAACAAGTGC ATCGGCTTCA GCGTCAACGA GTGCGTCTGA ATCGGCATCA	4020
ACGAGTGCGT CCGCTTCAGC AAGTACTAGC GCCTCAGCCT CAGCGTCAAC AAGTGCATCG	4080
GCTTCAGCAT CAACGAGTGC GTCCGCTTCA GCAAGTACTA GCGCCTCAGC CTCAGCGTCA	4140
ACAAGTGCGT CCGGTTTCAGC GTCAACGAGT GCGTCTGAGT CAGCATCAAC GAGTGCGTCA	4200
CCTCAKCAAG CACATCAGCT TCTGAATCTG CATCAACCAG TGCCTCACTT CCGCATCAAC	4260
AAGCGCCTCG GCCTCAGCAA GTACAAGTGC TTCAGCCTCA GCATCAACCA GTGCATCAGC	4320
TTCAGCCTCA ACAAGTGCTT CAGCCTCAGC GTCAGACCAG TGCCTCGGCT TCAGCAAGTA	4380
CCAGTGCGTC ACTTCAGCAA GCACAAGTGC GTCAGCTTCA GCATCAACCA GTGCTTCGGC	4440
TTCGGCATCA ACAAGTGCTT CAGCATCAGC ATCAACGAGT GCG	4483

(2) INFORMATION FOR SEQ ID NO: 364:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2550 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 364:

1379

GTACCTCAGC GTCCTTCCGC CTCAACCAGT GCGTCGCTT CAGCAAGCAC AAGTGCGTCA	60
CCTCAGCAAG TATCTCAGCG TCTGAATCGG CATCAACGAG TCGGTCGGCC TCAGCAAGCG	120
CAAGTACCTC AGCGTCACTT CCGCCTCAAC CAGTGCGTCG GCTTCAGCAA GCACAAGTGC	180
GTCAsCTCAG CAAGTATCTC AGCGTCTGAA TCGGCATCAA CGAGTGCGTC TGAGTCAGCA	240
TCAACGAGTA CGTCAGCCTC AGCAAGCACA TCAGCTTCTG AATCGGCATC AACCAGTGCG	300
TCAGCCTCAG CATCGACAAG CGCCTCAGCT TCAGCAAGTA CCAGTGCTTC AGCCTCAGCG	360
TCGACAAGTG CGTCGGCCTC AACCAGTGCA TCTGAATCGG CATCAACCAG TCGTCAGCC	420
TCAGCAAGTA CTAGTGATC AGCTTCAGCA TCAACGAGTG CATCGGCTTC AGCATCAACC	480
AGTGCCTCGG CTTTCAGCGTC AACCAGTGCG TCAGCTTCAG CAAGTACCAG TGCTTCAGTC	540
TCAGCATCAA CAAGTGCTTC AGCCTCAGCA TCGACAAGTG CCTCGGCTTC AGCAAGCACA	600
TCAGCATCTG AATCAGCGTC GACAAGTGCG TCGGCCTCAA CCAGTGATC TGAATCGGCA	660
TCAACCAGTG CGTCAGCCTC AGCAAGTACT AGTGATCAG CTTTCAGCATC AACGAGTGCA	720
TCGGCTTCGG CGTCAACCAG TGCATCAGAG TCAGCAAGTA CCAGTGCGTC AGCTTCCGCA	780
TCAACAAGTG CCTCGGCTTC AGCAAGCACA TCAGCATCTG AATCAGCGTC AACCAGTGCT	840
TCGGCTTCAG CAAGTACCAG TGCTTCAGCT TCAGCATCAA CCAGCGCCTC GGCCTCAGCA	900
AGCACCTCAG CTTCTGAATC GGCCTCAACC AGCGCCTCGG CCTCAGCAAG CACCTCAGCT	960
TCTGAATCGG CCTCAACCAG CGCCTCAGCC TCAGCATCAA CGAGTGCTTC GGCTTCAGCA	1020
AGCACAAGCG CCTCGGCTTC AGCATCAACG AGTACGTCAG CTTTCAGCGTC AACCAGTGCT	1080
TCAGCCTCAG CATCAACAAG TCGTCAGCC TCAGCAAGTA TCTCAGCGTC TGAATCGGCA	1140
TCAACGAGTG CGTCTGAGTC AGCATCAACG AGTACGTCAG CCTCAGCAAG CACAAGTGCT	1200
TCAGCCTCAG CAAGTATCTC AGCGTCTGAA TCGGCATCAA CGAGTGCGTC CGCTTCAGCA	1260
AGTACTAGCG CCTCAGCATC AGCGTCAACA AGTGCTTCGG CTTTCAGCGTC AACGAGTGCG	1320
TCTGAGTCAG CATCAACGAG TACGTAGCC TCAGCAAGCA CATCAGCTTC TGAATCTGCA	1380
TCAACCAGTG CGTCAGCCTC AGCATCGACA AGCGCCTCAG CTTTCAGCAAG TACCAGTGCG	1440
TCAGCCTCAG CAAGTACCAG TGCTTCAGCC TCAGCGTCGA CAAGTGCGTC GGCCTCAACC	1500
AGTGCACTG AATCGGCATC AACCAGTGCG TCAGCTCAGC AAGTACTAGT GCATCAGCTT	1560
CAGCATCAAC GAGTGATCG GCTTCGGCGT CAACCAGTGC ATCAGAGTCA GCAAGTACCA	1620
GTGCGTCACT TCCGCATCAA CAAGTGCTC GGCCTTCAGCA AGCACATCAG CATCTGAATC	1680
AGCGTCAACC AGTGCTTCGG CTTTCAGCAAG TACCAGTGCT TCAGCTTCAG CATCAACCAG	1740

1380

CGCCTCGGCC TCAGCAAGCA CCTCAGCTTC TGAATCGGCC TCAACCAGCG CCTCGGCCTC	1800
AGCAAGCACC TCAGCTTCTG AATCGGCCTC AACCAGCGCC TCAGCCTCAG CATCAACGAG	1860
TGCTTCGGCT TCAGCAAGCA CAAGCGCCTC GGGTTCAGCA TCAACGAGTA CGTCAGCTTC	1920
AGCGTCAACC AGTGCTTCAG CCTCAGCATC AACAAAGTGG TCAGCCTCAG CAAGTATCTC	1980
AGCGTCTGAA TCGGCATCAA CGAGTGCCTC TGAGTCAGCA TCAACGAGTA CGTCAGCCTC	2040
AGCAAGCACC TCAGCTTCTG AATCGGCCTC AACCAGTGG TCAGCCTCAG CATCGACAAG	2100
CGCCTCAGCT TCAGCAAGTA CCAGTGCTTC AGCCTCAGCG TCGACAAGTG CGTCGGCCTC	2160
AACCAGTGCA TCTGAATCGG CATCAACCAG TGGTCAGCC TCAGCAAGTA CTAGTGCATC	2220
GGCTTCAGCA TCAACCAGTG CCTCGGCTTC AGCGTCAACC AGTGCGTCAG CTTAGCAAG	2280
TACCAGTGCT TCAGTCTCAG CATCAACAAG TGCTTCAGCC TCAGCATCGA CAAGTGCCTC	2340
GGCTTCAGCA AGCACATCAG CATCTGAATC AGCGTCGACA AGCGCCTCAG CTTAGCAAG	2400
TACCAGTGCG TCAGCCTCAG CGTCGACAAG TGGTCAGCT ACAGCAAGTA CTAGTGCATC	2460
AGCTTCAGCA TCAACGAGTG CATCGGCTTC GCGTCAACC AGTGCATCAG AGTCAGCAAG	2520
TACCAGTGCG TCAGTTCACG CATCAACAAG	2550

(2) INFORMATION FOR SEQ ID NO: 365:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1436 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 365:

ACCCAGCAAG TACTAGTGCA TCGGCTTCAG CAAGCACCAG TGGTCGGCT TCAGCATCAA	60
CCAGTGCCTC AGCCTCAGCA AGTATCTCAG CGTCTGAATC GGCATCAACG AGTGCGTCAC	120
CTCAGCAAGT ACTAGTGCAT CAGCATCAGC ATCAACGAGT GCATCGGCTT CAGCAAGTAC	180
CAGCGCCTCA GCTTCAGCAA GCACCAGTGC GTCAsCTCAG CAAGTACCAG CGCCTCAGCC	240
TCAGCAAGCA CCAGTGCCTC AGCTTCAGCA AGTACCAGTG CGTCAGCCTC AGCGTCGACA	300
AGTGCGTCGG CTTAGCAAG TACCTCAGCG TCTGAATCAG CATCAACGAG TGCATCAGCT	360
TCAGCATCAA CAAGTGCTTC AGCTTCAGCA AGTATCTCAG CGTCTGAATC GGCATCAACG	420
AGTGCGTCCG CTTAGCAAG TACTAGCGCC TCAGCATCAG CGTCAACAAG TGCTTCGGCT	480
TCAGCGTCAA CGAGTGCCTC TGAGTCAGCA TCAACGAGTA CGTCAGCCTC AGCAAGCACA	540
TCAGCTTCTG AATCTGCATC AACCAGTGGC TCAGCCTCAG CATCGACAAG CGCCTCAGCT	600

1381

TCAGCAAGTA CCAGTGCCTC AGCCTCAGCA AGTACCAGTG CTTCAGCCTC AGCGTCGACA	660
AGTGCGTCGG CCTCAACCAG TGCATCTGAA TCGGCATCAA CCAGTGCCTC AGCCTCAGCA	720
AGTACTAGCG CCTCAGCCTC AGCATCAACG AGTGCGTCCG CTTCAGCAAG TACTAGTGCA	780
TCAGCTTCAG CAAGTACTAG CGCCTCAGCC TCAGCGTCGA CAAGCGCCTC AGCTTCAGCA	840
AGTACCAGTG CGTCAGCCTC AGCGTCGACA AGTGCGTCGG CTTCAGCAAG TACCTCAGCG	900
TCTGAATCAG CATCAACAAG TGCCTCGGCT TCAGCATCAA CGAGTGCATC AGCTTCAGCA	960
TCAACAAGTG CTTCAGCTTC AGCAAGTACC AGTGCGTCGG CTTCAGCATC AACGAGTGCT	1020
TCAGTCTCAG CGTCAACCAG TGCCTCTGAA TCCGCATCAA CAAGTGCCTC GGCTTCAGCA	1080
AGCACCAGTG CTTGGCTTC AGCGTCAACG AGTGCGTCTG AGTCAGCATC AACGAGTGCG	1140
TCAGCCTCAG CAAGCACATC AGCTTCTGAA TCTGCATCAA CCAGTGCCTC AGCTTCCGCA	1200
TCAACAAGCG CCTCGGCCTC AGCAAGTACA AGTGCTTCAG CTTCAGCATC AACCAGTGCA	1260
TCAGCTTCAG CCTCAACAAG TGCTTCAGCC TCAGCGTCAA CCAGTGCCTC GGCTTCAGCA	1320
AGTACCAGTG CGTCAGCTTC AGCAAGCACA AGTGCGTCAG CTTCAGCATC AACCAGTGCT	1380
TCGGCTTCGG CATCAACAAG TGCCTCAGCA TCAGCATCAA CGAGTGCCTC AGCCGG	1436

(2) INFORMATION FOR SEQ ID NO: 366:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 735 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 366:

GCAGTTGCCA CACCGTGCTG ACCAGCACCC GTTCCTGCGA TAATTTTCTT TTTACCCATG	60
CGTwTGGCAA GCCAAACTTG TCCTAAGGCA TTGTAAATCT TGTGGGCTCC TGATGGTTA	120
AGGTCTTCCC GTTTGAGATA AATCTTGCTC CGCCAATATG CTGGGTCAAG TTTTTCGCT	180
AATAAAGAGG AGTTTCAGT CCTACGTACT GCGCAAAAG CTGGTTAAT TCCTCTTGGA	240
AACTTGGGTC TGCCTGACTT TCACGGTAGG CTTCTCCAA CTCCAAACT GCTGTATCA	300
ATGTTTCTGG GACAAAACGT CCGCCGAATT TTCCGTAAAA TCCATCTTTA TTTGGTTCCT	360
GATATGCCAT GCTTTACCCT CTCTATAAAT CTTCTAATCT TTTCATGATC TTTTGTCCA	420
TCTGTCTCCA CTCCGCTCGA TACATCTACT GCATAGGGAG TAAAGTGTG AATTGCTTTT	480
ACTACATTAT CTTCAATAAG GCCACCTGCG ATAAAGAAGG GCTGTGCTAG TCCAGTCGTA	540

1382

TCCAGTTGAC CCCAATCAAA GGGCTGGCCA CTTCTTGCCA CAGGGGCATC AAAGAGTAGA	600
TAATCTGCCT GAGAATTGGG GACATGCCCC TTTCATCTA CCTGCACAGC CTGAATACTG	660
GCACAAGGCA AATTCTCAAA TAAATCATCT GCCACCTGAC CGTGAACCTG AACCAAGTCC	720
AAGCCGGGGA TCCTC	735

(2) INFORMATION FOR SEQ ID NO: 367:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1702 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 367:

TACTAGCGCC TCAGCCTCAG CGTCAACAAG TGCATCGGCT TCAGCATCAA CGAGTGCCTC	60
CGCTTCAGCA AGTACTAGCG CCTCAGCCTC AGCGTCAACA AGTGCATCGG CTTCAGCGTC	120
AACGAGTGC G TCTGAGTCAG CATCAACGAG TGGCTCAGCC TCAGCAAGCA CATCAGCTTC	180
TGAATCTGCA TCAACCAAGTG CGTCAGCCTC AGCATCGACA AGCGCCTCAG CTTCAGCAAG	240
TACCAAGTGC TCAGCCTCAG CGTCGACAAG TGGCTCGGCT TCAGCAAGTA CCAGTGCCTC	300
AGCCTCAGCA AGTACCAAGTG CGTCAGCCTC AGCGTCGACA AGTGCCTCGG CCTCAACCAG	360
TGCATCTGAA TCGGCATCAA CCAGTGCCTC AGCCTCAGCA AGTACTAGTG CATCAGCTTC	420
AGCATCAACG AGTGCATCGG CTTCAGCATC AACCAAGTGA TCAGAGTCAG CAAGTACCAG	480
TGGCTCAGCT TCCGCATCAA CAAGTGCCTC GGCTTCAGCA AGTACTAGCG CCTCAGCCTC	540
AGCGTCAACA AGTGCTTCAG CTTCCGCGTC AACCAAGCGC TCGGCCTCAG CAAGTATCTC	600
AGCGTCTGAA TCGGCATCAA CAAGTGCCTC GGCTTCAGCA TCAACGAGTG CATCAGTCTC	660
AGCAAGCACC AGTGCCTCGG CCTCAGCAAG CACCAGCGCG TCTGAATCCG CATCAACCAG	720
TGCCTCAGCT TCAGCAAGTA CCTCAGCATC TGAATCAGCA TCAACAAGTG CATCGGCTTC	780
AGCAAGCACA AGTGCTTCAG CCTCAGCAAG TATCTCAGCG TCTGAATCCG CATCAACGAG	840
TGGCTCCGCT TCAGCAAGTA CTAGCGCCTC AGCATCAGCG TCAACAAGTG CTTCCGCTTC	900
AGCGTCAACG AGTGCCTCTG AGTCAGCATC AACGAGTACG TCAGCCTCAG CAAGCACATC	960
AGCTTCTGAA TCTGCATCAA CCAGTGCCTC AGCCTCAGCA TCGACAAGCG CCTCAGCTTC	1020
AGCAAGTACC AGTGCCTCAG CCTCAGCAAG TACCAGTGCT TCAGCCTCAG CGTCGACAAG	1080
TGGCTCGGCC TCAACCAAGTG CATCTGAATC GGCATCAACC AGTGCCTCAG CCTCAGCAAG	1140
TACTAGCGCC TCAGCCTCAG CATCAACGAG TGGCTCCGCT TCAGCAAGTA CTAGTGCATC	1200

1383

AGCATCAGCA TCAACGAGTG CATCGGCTTC AGCAAGTACC AGCGCCTCAG CTTCAGCAAG	1260
CACCAGTGCG TCAGCCTCAG CAAGTACCAG CGCCTCAGCC TCAGCAAGCA CCAGTGCCTC	1320
AGCTTCAGCA AGTACCAGTG CGTCAGCCTC AGCGTCGACA AGTGCGTCGG CTTCAGCAAG	1380
TACCTCAGCG TCTGAATCAG CATCAACGAG TGCATCAGCT TCAGCATCAA CAAGTGCTTC	1440
AGCTTCAGCA AGTACCAGTG CGTCGGCTTC AGCATCAACG AGTGCTTCAG TCTCAGCGTC	1500
AACCAGTGCC TCTGAATCAG CATCAACAAG TGCTCGGCT TCAGCAAGCA CCAGTGCCTC	1560
GGCTTCAGCA AGTACTAGTG CATCGGCTTC AGCATCGACA AGTGCGTCTG AATCGGCATC	1620
AACGAGTGCT TCGGCTTCAG CATCAACGAG TGCGTCAGCC TCAGCAAGCA CATCAGCTTC	1680
TGAATCTGCA TCAACCAGTG CG	1702

(2) INFORMATION FOR SEQ ID NO: 368:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 941 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 368:

ACCAGTGCAT CAGCTTCAGC CTCAACAAGT GCTTCAGCCT CAGCGTCAAC CAGTGCCTCG	60
GCTTCAGCAA GTACCAGTGC GTCACTTCAG CAAGCACAAG TGCGTCACTT CAGCATCAAC	120
CAGTGCTTCG GCTTCGGCAT CAACAAGTGC CTCAGCATCA GCATCAACGA GTGCGTCACC	180
TCAGCAAGTA CTAGTGCATC AGCATCAGCA TCAACCAGTG CATCAGCCTC AGCAAGTATC	240
TCAGCGTCTG AATCGGCATC AACGAGTGCA TCAGCATCAG CATCAACGAG TGCATCGGCT	300
TCAGCGTCAA CCAGTGCATC AGTCTCAGCA AGCACCAGTG CGTCGGCTTC AGCATCAACG	360
AGTGCCTCAG CCTCAGCAAG TATCTCAGCG TCTGAATCGG CATCAACGAG TCGTCAGCC	420
TCAGCAAGTA CTAGTGCATC GGCTTCAGCA AGCACCAGTG CGTCGGCTTC AGCATCAACC	480
AGTGCCTCAG CCTCAGCAAG TATCTCAGCG TCTGAATCGG CATCAACGAG TCGTCAGCC	540
TCAGCAAGTA CTAGTGCATC AGCATCAGCA TCAACGAGTG CATCGGCTTC AGCAAGTACC	600
AGCGCCTCAG CTTCAGCAAG CACCAGTGCG TCAGCCTCAG CAAGTACCAG CGCCTCAGCC	660
TCAGCAAGCA CCAGTGCCTC AGCTTCAGCA AGTACCAGTG CGTCAGCCTC AGCGTCGACA	720
AGTGCGTCGG CTTCAGCAAG TACCTCAGCG TCTGAATCAG CATCAACGAG TGCATCAGCT	780
TCAGCATCAA CAAGTGCTTC AGCTTCAGCA AGTACCAGTG CGTCGGCTTC AGCATCAACG	840

1384
AGTGCTTCAG TCTCAGCGTC AACCAGTGCC TCTGAATCAG CATCAACAAG TGCCTCGGCT 900
TCAGCAAGCA CCAGTGCGTC GGCTTCAGCA AGTACTAGTG C 941

(2) INFORMATION FOR SEQ ID NO: 369:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 869 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 369:

CAGCAAGTAC TAGTGATCA GCTTCAGCAT CAACGAGTGC ATCGGCTTCT GCGTCAACCA 60
GTGCATCAGA GTCAGCAAGT ACCAGTGCGT CAGCTTCCGC ATCAACAAGT GCCTCGGCTT 120
CAGCAAGCAC CAGTGCGTCG GCTTCAGCAA GTACTAGCGC CTCAGCCTCA GCCTCAACCA 180
GTGCGTCAGC CTCAGCAAGT ATCTCAGCGT CTGAATCGGC ATCAACGAGT GCGTCCGCTT 240
CAGCAAGTAC TAGCGCCTCA GCCTCAGCGT CAACAAGTGC ATCGGCTTCA GCGTCAACGA 300
GTGCGTCTGA ATCGGCATCA ACGAGTGCGT CCGCTTCAGC AAGTACTAGC GCCTCAGCCT 360
CAGCGTCAAC AAGTGATCG GCTTCAGCAT CAACGAGTGC GTCCGCTTCA GCAAGTACTA 420
GCGCCTCAGC CTCAGCGTCA ACAAGTGCGT CGGCTTCAGC GTCAACGAGT GCGTCTGAGT 480
CAGCATCAAC GAGTGCGTCA GCCTCAGCAA GCACATCAGC TTCTGAATCT GCATCAACCA 540
GTGCGTCAGC CTCAGCATCG ACAAGCGCCT CAGCTTCAGC AAGTACCAGT GCGTCAGCCT 600
CAGCGTCGAC AAGTGCGTCG GCTTCAGCAA GTACCAAGTGC GTCAGCCTCA GCAAGTACCA 660
GTGCGTCAGC CTCAGCGTCG ACAAGTGCGT CGGCCTCAAC CAGTGATCTT GAATCGGCAT 720
CAACCAAGTGC GTCAGCCTCA GCAAGTACTA GTGCATCAGC TTCAGCATCA ACGAGTGCGT 780
CGGCTTCAGC ATCAACCAAGT GCATCAGAGT CAGCAAGTAC CAGTGCGTCA GntTCCGCAT 840
GCAACAAGTG CCTCGGCTTC AGCAAGTAC 869

(2) INFORMATION FOR SEQ ID NO: 370:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 750 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 370:

TCAACAAGTG CCTCAGCATC AGCATCAACG AGTGCGTCAG CCTCAGCAAG TACTAGTGCA 60

1385

TCAGCATCAG CATCAACCAG TGCATCAGCC TCAGCAAGTA TCTCAGCGTC TGAATCGGCA	120
TCAACGAGTG CATCAGCATC AGCATCAACG AGTGCATCGG CTTCAGCGTC AACCAGTGCA	180
TCAGTCTCAG CAAGCACCAG TCGCTCGGCT TCAGCATCAA CGAGTGCCTC AGCCTCAGCA	240
AGTATCTCAG CGTCTGAATC GGCATCAACG AGTGCCTCAG CCTCAGCAAG TACTAGTGCA	300
TCGGCTTCAG CAAGCACCAG TCGCTCGGCT TCAGCATCAA CCAGTGCCTC AGCCTCAGCA	360
AGTATCTCAG CGTCTGAATC GGCATCAACG AGTGCCTCAG CCTCAGCAAG TACTAGTGCA	420
TCAGCATCAG CATCAACGAG TGCATCGGCT TCAGCAAGTA CCAGCGCCTC AGCTTCAGCA	480
AGCACCAGTG CGTCAGCCTC AGCAAGTACC AGCGCCTCAG CCTCAGCAAG CACCAGTGCC	540
TCAGCTTCAG CAAGTACCAG TCGCTCAGCC TCAGCGTCGA CAAGTGCCTC GGCTTCAGCA	600
AGTACCTCAG CGTCTGAATC AGCATCAACG AGTGCATCAG CTTCAGCATC AACAAGTGCT	660
TCAGCTTCAG CAAGTATCTC AGCGTCTGAA TCGGCATCAA CGAGTGCCTC CGCTTCAGCA	720
AGTACTAGCG CCTCAGCATC AGCGTCAACG	750

(2) INFORMATION FOR SEQ ID NO: 371:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 957 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 371:

CCGGAAAACA GCTCTGGCGC TTGGTCTTGC CCAGCGTATT GCTAGTGGTG ACGTGCCTGC	60
GGAAATGGCT AAGATGCGCG TGTTAGAACT TGATTGATG AATGTCGTG CAGGGACACG	120
CTTCCGTGGT GACTTTGAAG AACGCATGAA TAATATCATC AAGGATATTG AAGAAGATGG	180
CCAAGTCATC CTCTTTATCG ATGAACTCCA CACCATCATG GGTTCCTGGTA GCGGGATTGA	240
TTGACTCTG GATGCGGCCA ATATCTTGAA ACCAGCCTTG GCGCGTGGAA CTTTGAGAAC	300
GGTTGGTGCC ACTACTCAGG AAGAATATCA AAAACATATC GAAAAAGATG CGGCACTTTC	360
TCGTCGTTTC GCTAAAGTGA CGATTGAAGA ACCAAGTGTG GCAGATAGTA TGA TATTTT	420
ACAAGGTTTG AAGGCGACTT ATGAGAAACA TCACCGTGTA CAAATCACAG ATGAAGCGGT	480
TGAAACAGCG GTTAAGATGG CTCATCGTTA TTAAACCACT CGTCACTTGC CAGACTCTGC	540
TATCGATCTC TTGGATGAGG CGGCAGCAAC AGTGCAAAAT AAGGCAAAGC ATGTAAAAGC	600
AGACGATTCA GATTTGAGTC CAGCTGACAA GGCCCTGATG GATGGCAAGT GGAAACAGGC	660

1386

AGCCCAGCTA ATCGCAAAAG AAGAGGAAGT ACCTGTCTAC AAAGACTTGG TGACAGAGTC	720
TGATATTTTG ACCACCTTGA GTCGCTTGTC AGGAATCCCA GTTCAAAAAC TGACTCAAAC	780
GGATGCTAAG AAGTATTTAA ATCTTGAAGC AGAACTCCAT AAACGGGTTA TCGGTCAAGA	840
TCAAGCTGTT TCAAGCATTG GCCGTGCCAT TCGCCGCAAC CAGTCAGGGA TTCGCAGTCA	900
TAAGCGTCCG ATTGGTTCCT TTATGTTTCT AGGGCCTACA GGTGTCGGGG TATCCGA	957

(2) INFORMATION FOR SEQ ID NO: 372:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 807 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 372:

CAAAGCGCCT CAGCTTCAGC ATCAACAAGT GCGTCGGCTT CAGCATCAAC CAGTGCCTCG	60
GCTTCAGCGT CAACCAAGTGC GTCACATTCA GCAAGTACCA GTGCTTCAGT CTCAGCATCA	120
ACAAGTGCTT CAGCCTCAGC ATCGACAAGT GCCTCGGCTT CAGCAAGCAC ATCAGCATCT	180
GAATCAGCGT CAACCAAGTGC TTCGGCTTCA GCAAGTACCA GTGCTTCAGC TTCAGCATCA	240
ACCAGCGCCT CGGCTTCAGC AAGCACCCTCA GCTTCTGAAT CGGCTTCAAC CAGCGCCTCG	300
GCCTCAGCAA GCACCTCAGC TTCTGAATCG GCCTCAACCA GCGCTCAGC CTCAGCATCA	360
ACGAGTGCTT CGGCTTCAGC AAGCACAAGC GCCTCGGGTT CAGCATCAAC GAGTACGTCA	420
GCTTCAGCGT CAACCAAGTGC TTCAGCCTCA GCATCAACAA GTGCGTCAGC CTCAGCAAGT	480
ATCTCAGCGT CTGAATCGGC ATCAACGAGT GCGTCTGAGT CAGCATCAAC GAGTACGTCA	540
GCCTCAGCAA GCACCTCAGC TTCTGAATCG GCCTCAACCA GTGCGTCAGC CTCAGCATCG	600
ACAAGCGCCT CAGCTTCAGC AAGTACCAGT GCTTCAGCCT CAGCGTCGAC AAGTGCCTCG	660
GCCTCAACCA GTGCATCTGA ATCGGCATCA ACCAGTGCCT CAGCCTCAGC AAGTACTAGT	720
GCATCGGCTT CAGCATCAAC CAGTGCCTCG GCTTCAGCGT CAACCAAGTGC GTCAGCTTCA	780
GCAAGTACCA TGTGCTTCAT GTCTCAG	807

(2) INFORMATION FOR SEQ ID NO: 373:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1068 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

1387

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 373:

CATCGGCTTC AGCATCAACG AGTGCCTCCG CTTAGCAAG TACTACCGCC TCAGCCTCAG	60
CGTCAACAAG TGCATCGGCT TCAGCGTCAA CGAGTGCCTC TGAGTCAGCA TCAACGAGTG	120
CGTCACCTCA GCAAGCACAT CAGCTTCTGA ATCTGCATCA ACCAGTGCCT CACCTCAGCA	180
TCGACAAGCG CCTCAGCTTC AGCAAGTACC AGTGCCTCAC CTCAGCGTCG ACAAGTGCCT	240
CGGCTTCAGC AAGTACCAGT GCGTCAsCTC AGCAAGTACC AGTGCCTCAC CTCAGCGTCG	300
ACAAGTGCCT CGGCCTCAAC CAGTGCATCT GAATCGGCAT CAACCAAGTGC GTCACCTCAG	360
CAAGTACTAG TGCATCAGCT TCAGCATCAA CGAGTGCATC GGCTTCAGCA TCAACCAAGTG	420
CATCAGAGTC AGCAAGTACC AGTGCCTCAG cTTCCGCATC AACAAGTGCC TCGGCTTCAG	480
CAAGTACTAG CGCCTCAGCC TCAGCGTCAA CAAGTGCTTC AGCTTCCGCG TCAACCAAGCG	540
CCTCGGCCTC AGCAAGTATC TCAGCGTCTG AATCGGCATC AACAAGTGCC TCGGCTTCAG	600
CATCAACGAG TGCATCAGTC TCAGCAAGCA CCAAGTGCCTC GGCCTCAGCA AGCACCAGCG	660
CGTCTGAATC CGCATCAACC AGTGCCTCAG CTTAGCAAG TACTAGCGCC TCAGCATCAG	720
CATCAACAAG TGCATCGGCT TCAGCAAGCA CAAGTGCTTC AGCCTCAGCA AGTATCTCAG	780
CGTCTGAATC GGCATCAACG AGTGCCTCCG CTTAGCAAG TACTAGCGCC TCAGCATCAG	840
CGTCAACAAG TGCTTCGGCT TCAGCGTCAA CGAGTGCCTC TGAGTCAGCA TCAACGAGTA	900
CGTCAGCCTC AGCAAGCACA TCAGCTTCTG AATCTGCATC AACCAGTGCG TCAGCCTCAG	960
CATCGACAAG CGCCTCAGCT TCAGCAAGTA CCAAGTGCCTC AGCCTCAGCA AGTACCAGTG	1020
CTTCAGCCTC AGCGTCGACA AGTGCCTCGG GCTCAACCAG TGCACTCTG	1068

(2) INFORMATION FOR SEQ ID NO: 374:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 620 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 374:

CAGCATCAAC GAGTGCTTCA GTTTCAGCGT CAACCAAGTGC CTCTGAATCA GCTTCAACAA	60
GTGCCTCGGC TTCAGCAAGC CCCAGTGCCT CGGCTTCAGC AAGTACTAGT GCATCGGCTT	120
CAGCATCGAC AAGTGCCTCT GAATCGGCAT CAACGAGTGC TTCGGCTTCA GCATCAACGA	180
GTGCGTCAGC CTCAGCAAGC ACATCAGCTT CTGAATCTGC ATCAACCAGT GCGTCCGyTT	240

1388

CAGCGTCAAC CAGTGCCTCG GCTTCAGCGT CGACAAGTGC TTCGGCTTCA GCATCAACGA	300
GTGCGTCGGC CTCAGCAAGC GCAAGTACCT CAGCGTCAGC TTCCGCTTCA ACCAGTGCCT	360
CGGCTTCAGC AAGCACAAGT GCGTCAGCCT CAGCAAGTAT CTCAGCGTCT GAATCGGCAT	420
CAACGAGTGC GTCTGAGTCA GCATCAACGA GTACGTCAGC CTCAGCAAGC ACATCAGCTT	480
CTGAATCTGC ATCAACCAGT GCGTCAGCCT CAGCATCGAC AAGCGCTTCA GCTTCAGCAA	540
GTACCACTGC TTCAGCTTCA GCGTCGACAA GTGCGTCGGC CTCAACCAGT GCATCTGAAT	600
CGGCATCAAC CAGTGCCTCA	620

(2) INFORMATION FOR SEQ ID NO: 375:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 720 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 375:

GTATTGGGGC GCCCCAACCT CTATGTGACT ACGGATTATT TCCTAGATTA CATGgGGATA	60
AACCATTTAG AAGAATTACC AGTGATTGAT GAGCTTGAGA TTCAAGCCCA AGAAAGCCAA	120
TTATTTGGTG AAAGGATAGA AGAAGATGAG AATCAATAAG TATATTGCCC ACGCAGGTGT	180
GGCCAGTAGG AGAAAAGCAG AAGAGCTGAT TAAGCAAGGC TTGGTGACGG TTAACGGCCA	240
AGTGGTGCCT GAACTAGCAA CCACTATCAA GTCAGGCGAC AAGGTCGAAG TTGAAGGTCA	300
ACCTATCTAC AACGAAGAAA AGGTCTACTA TCTGCTTAAC AAACCACGCG GTGTGATTTT	360
CAGTGTGACA GATGATAAGG GTCGCAAGAC GGTGTGCGAC CTCTTGCCCA ATGTCAAAGA	420
GCGTATTTAC CCTGTGGGTC GTTTGGACTG GGATACATCA GGTGTCTTGA TTTTGACCAA	480
TGATGGGGAC TTTACAGACG AGATGATTCA CCCTCGTAAT GAGATTGACA AGGTTTATGT	540
CGCGCGTGTT AAAGGTGTGG CCAATAAGGA CAATCTCCGC CCCTTGACCC GTGGTCTTGA	600
GATTGATGGT AAGAAAACCA AGCCATAATA TATAGGTTTT GTAGCCTCTA CACCATAAAT	660
ATTTGCTAAT AAAAATACTG TATTATTACC CTCTTAAGGT GCGAAATTAT TCAAGTTCTT	720

(2) INFORMATION FOR SEQ ID NO: 376:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 648 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

1389

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 376:

CGCCATTTC	CATCGTACCG	CCGAAAATCC	CAGCGCCTCA	GCCATCAAAT	ATCCTATCAA	60
CGTTCTCAA	AAAAGTGACC	GCTCTCTCAT	CATGTTTCCA	AGTGGTAGCC	GCCACTCAA	120
CGATGTCAAG	GGGGCGCAC	ACTskATTGC	CAAAATGGCC	AAGGTCCGTA	TCATGCCGGT	180
TACCTACACC	GGTCCCATGA	CTTTGAAGGG	CTTGATTAGC	CGTGAACGTG	TCGATATGAA	240
CTTTGGAAAT	CCAATCGATA	TCTCAGATAT	CAAGAAAATG	AATGATGAAG	GCATTGAAAC	300
AGTCGCCAAT	CGTATTCAA	CAGAATTCCA	ACGTCTGGAC	GAAGAAACGA	AACAATGGCA	360
CAATGATAAA	AAACCAAATC	CACTCTGGTG	GTTTATCCGC	ATCCCTGCCC	TCATCCTTGC	420
TATTATCCTC	GCTATCCTAA	CCATCATCTT	TAGCTTTATC	GCAAGCTTCA	TCTGGAACCC	480
AGATAAGAAA	AGAGAAGAAC	TTGCATAGAA	GAAATGAACC	TTGGCCAAAC	AGCTAAGGTT	540
TTCATTTATA	TAGTAGATTG	GwACTAGAAT	AGTACACCTC	TACTTCTAAA	ACATTTTATG	600
AAATCGATTT	GACTGTCCTG	ATCGATTTGT	CCTAATCTTA	TTTCAATT		648

(2) INFORMATION FOR SEQ ID NO: 377:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 690 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 377:

GTGCATCGCT	TTCAGCATCG	ACAAGTGGT	CTGAATCGGC	ATCAACGAGT	GCTTCGGCTT	60
CAGCATCAAC	GAGTGCGTCA	GCTTCAGCAA	GCACATCAGC	TTCTGAATCT	GCATCAACCA	120
GTGCGTCCGC	TTCAGCGTCA	ACCAGTGGT	CGGCTTCAGC	GTGACAAGT	GCTTCGGCTT	180
CAGCATCAAC	GAGTGCGTCG	GCCTCAGCAA	GCGCAAGTAC	CTCAGCGTCA	GCTTCCGCCT	240
CAACCAAGTGC	GTCCGCTTCA	GCAAGCACA	GTGCGTCAGC	CTCAGCAAGT	ATCTCAGCGT	300
CTGAATCGGC	ATCAACGAGT	GCGTCGGCCT	CAGCAAGCGC	AAGTACCTCA	GCGTCAGCTT	360
CCGCCTCAAC	CAGTGCGTCG	GCTTCAGCAA	GCACAAGTGC	GTGAGCCTCA	GCAAGTATCT	420
CAGCGTCTGA	ATCGGCATCA	ACGAGTGGT	CTGAGTCAGC	ATCAACGAGT	ACGTCAGCCT	480
CAGCAAGCAC	ATCAGCTTCT	GAATCGGCAT	CAACCAAGTGC	GTGAGCCTCA	GCATCGACAA	540
GCGCCTCAGC	TTCAGCAAGT	ACCAGTGGT	CAGCCTCAGC	GTGACAAGT	GCGTCGGCCT	600
CAACCAAGTGC	ATCTGAATCG	GCATCAACCA	GTGCGTCAGC	CTCAGCAAGT	ACTAGTGCAT	660

CAGCTTCAGC ATCAACGAGT GCATCGGCTT 1390 690

(2) INFORMATION FOR SEQ ID NO: 378:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1003 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 378:

CGAGATTCTC TGGAGTTATG GATGTCGTTC CAATATGTGC ACGTTGGAAT GTTAGTGCTT 60
ATATGGGGGG AACAGAATCC TCTCTTGATT GAAGACAAGC TAGTCATTAG GCTGGTTTGT 120
CTTTTGTCA ACTGTAGTGG GTTGATATAA TAGTATTAGT GAGTGGGATA AAAGTTTCAT 180
TTAGTTTATT CAGTACAAAT TTAACGGGTC AAGATTATA TACTAGTGGT GTTTTGGGG 240
CTGAGAGAAG TATCTTGATT TTATGTGTGG TTTTATACT TACAGTTGTT CTGCTCCAAA 300
GAGCTTGTAG AGAAGAATTA GCTCATAAAG GAGATTGATT ATTTTGATAT CAAAAAATG 360
CACAGATAA CCTGATGCAT TTTTTCGCG ACAATGCTTG CTACTTCCTT CTGTCGAATT 420
TAGACAATTT TAAACCCCAA TTATTCACCC CAAATCTAAA AACCATCCAG AATCCTTGCC 480
TTAGCTTAGA TCCTGGATGG TTTCTTTTTT CACCCAATGG GTGTTTTTTA CTAGACAAAA 540
AAGAGTTTCC CCTTTATGGT ATAAGTGTAG AAAAAACAC AAAAAGAAAG GAAACTCACA 600
TGAACAGTTT ACCAAATCAT CACTTCCAAA ACAAGTCTTT TTACCAACTA TCTTTCGATG 660
GAGGTCATTT AACCAGTAT GGTGGTCTTA TCTTTTTTCA GGAACTTTTT TCCCAGTTGA 720
AACTAAAAGA GCGGATTCTT AAGTATTTAG TAACGAATGA CCAACGCCGC TACTGTCGTT 780
ATTGCGGATTC AGATATCCTT GTCCAGTTCC TCTTCAACT GTTAACAGGT TATGGAACGG 840
ACTATGCTTG TAAAGAATTG TCAGCTGATG CCTACTTTCC AAAATTATTG GAAGGAGGGC 900
AGCTTGCTTC ACAGCCAACC TTATCCCGTT TTCTTCCAG AACTGACGAG GAAACAGTCC 960
ATAGTTTGCG ATGCCTCAAC CTTGAATTGG TCGAATTCTT TTT 1003

(2) INFORMATION FOR SEQ ID NO: 379:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 738 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 379:

1391

CCGATGATTC TGATTGGTTT GCTCTTTACT TTGCTGGGAA TTTTGAGGTA GATCTATGAT 60
TGAAATACTA ATTGTTTTAG CTATTATCCT ATCTCTTGCT TTGATTGTAT TGGTAACTAT 120
ACAACCCCGT CAAATCAAC TATTTTCCAT GGATGCCACT AGTAATATTG GTAAACCAAG 180
CTACTGGCAG AGCAACACCT TGGTCAAGGT GCTCACTTTA TTGGTGAGTT TGGCTTTATT 240
TATTCTACTA TTAACCTTTA TGGTGATTAC TTATAAATAA AAGAAAACCT CAGATATTCA 300
CCTTTTGTGG ATTGGTCTGA AGTTTCTTTT TTTATACTCA ATGAAAATCA AAGAGCAAAC 360
TAGGAAGCTA GCCGCACkGC TCAAAACACC GTTTTGAGGT TGTAGATATA ACTGACGAGc 420
GACTCAAAAC ACCGTTTTGA GGTGTAGAT ATAAC TGACG AGcGACTCAA AACACCGTTT 480
TGAGGTTGTG GATAGAACTG ACGAGcGACT CAAAACACCG TTTTGAGGTT GTGGATAGAA 540
CTGACGAAGT CGcTCAAAAC ACCGTTTTGA GGTGTGGAT AGAACTGACG AAtgctCAAA 600
ACACCGTTTT GAGGTTGTGG ATAGAACTGA CGAAGCgaaC ATATATACAG CAAGGCGACG 660
CTGACGTGGT TTGAAGAGTA TTACTGTCTA TATTTTGGT AAAAATCAAC TTTTACTTGG 720
ATGAAGGTTT TTTTTTTT 738

(2) INFORMATION FOR SEQ ID NO: 380:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 695 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 380:

CCGTCTTATC AAAGAGGTTA ACAAAGGCAC CAAATTTCTC GATACGAACG ACTTTAGCAC 60
GGTAAACTTC ATCCACTTTG GCTTCACGAA CCAAACCAGC AATAATTCTT TTGGCACGGT 120
TAATAGCATC TTGGTCACTA GAGTAGATAG ACACATTTC TCTTCGTCT ATATCAATCT 180
TAACACCTGT TTCAGCGATA ATCTTGTCGA TGGTTCTCC ACCCTTACCG ATGACAATCT 240
TAATCTTGTC CACATCAATC TTGATCGTAT CAATTTTCGG AGCAGTTGGA GCCAATTCTG 300
GACGAACTTC TGGAAATGGT GCTTCAATGA CATCAAGGAT TTCAAAACGC GCTTCTTGG 360
CTTGAGCAAG AGCCTCCGTC AAGATTTCTG CAGTAATCCC TTGAATCTTG ATATCCATTT 420
GAAGGGCTGT AATCCCATCA CGAGTACCTG CAACCTTGAA GTCCATATCT CCAAAGTGAT 480
CTTCCAAACC TTGGATATCT GTCAATACTG TGTAGTTATT TCCATCTGAG ATAAGCCCCA 540
TAGCAATACC AGCTACTGGC GCCTTGATTG GCACACCACC AGCCATAAGG GCAAGAGTTC 600

1392
CCGCACAGAT AGAAGCTTGA GATGAAGAAC CGTTTGATTC CAAAACCTTCT GCTACTAGAC 660
GGATAGCGTA GGGGAATTCT TCCAAGCTTG GCAGG 695

(2) INFORMATION FOR SEQ ID NO: 381:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 691 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 381:

GACATCTTAT CTAAATACAT GCTAATATAT TTAGATACAA ACATTCCAAC TTGATAATTT 60
TCACTCATCT TTCATCATTC CTTATACAAC TATGCAGTAT AAATAGAATA GTTTTCTCAT 120
CAGAATGAGA CTATTTTAAT ATTAGATCCC CAATTATTCA CCCCAAATCT AAAAACCATC 180
CAGAATCCCT GCCTTAGCTT AGATCCTGGA TGGTTTCTTT TTTCACCCAA TGGGTGTTTT 240
TTACTAGACA AAAAAGAGTT TCCCCTTTAT GGTATAAGTG TAGAAAAAAA CACAAAAAGA 300
AAGGAACTC ACATGAACAG TTTACCAAAT CATCACTTCC AAAACAAGTC TTTTACCAA 360
CTATCTTTTCG ATGGAGGTCA TTTAACCAG TATGGTGGTC TTATCTTTTT TCAGGAACTT 420
TTTTCCCACT TGAACTAAA AGAGCGGATT TCTAAGTATT TAGTAACGAA TGACCAACGC 480
CGCTACTGTC GTTATTCGGA TTCAGATATC CTTGTCCAGT TCCTCTTTCA ACTGTTAACA 540
GGTTATGGAA CGGACTATGC TTGTAAAGAA TTGTCAGCTG ATGCCTACTT TCCAAAATTG 600
TTGGAAGGAG GGCAGCTTGc TTCACAGCCA ACCTTATCCC GwTTTCTTTC CAGAACTGAC 660
GAGGAAACAG TCCATAGTTT GCGATGCCTC A 691

(2) INFORMATION FOR SEQ ID NO: 382:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 750 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 382:

ATCTCTCTGC GTAATGGTCC TCAGATAACT CTGATGATGT GTGGCGATAT AGAACTGAGC 60
CAAGTTATGC CTAAAGGGCC TTAGGAATAG GAGCTTTCAC AAGCTTATCC AGATGATTAT 120
CTTTTACTCG TTATGGACAA TGCTATATGG CATAAATCAA GTACCTTAAA GATTCCGACT 180
AATATTGGCT TTGCATTTAT TCCTCCATAC ACACCAGAGA TGAACCCCAT TGAACAAGTG 240

1393

TGGAAGAGA TTCGTAAACG TGGATTTAAG AATAAAGCCT TTCGAACCTT GGAAGATGTC	300
ATACAAGGAC TGGAGAAGGA GGTGATAAAG TCCATCGTTA ATCGGAGACG GACTAGAATG	360
CTTTTGTAAA ACAGATGAGT ATAAAAAGAA AGTCCTCATT TCAATAGAAA TCACGACTTT	420
CTGATGAATT TATAGTAAAA TGAAATAAGA ACAGGATAGT CAAATCGATT TCTAACAATG	480
TTTGTAGAAGC AGAGGTGTAC TATTCTAGTT TAAATCCACT ATATTGGGG AGTGATAGAA	540
AAGCCCTTCA TCAGCCAATC TACTTGTTCA GGTGCGAGAG CTTTGACATC CTTTCTGTA	600
CTGGACCAAG TCAGTTTCC GTTCTCAAAG CGTTTATATA ATATCCAAA TCCTTGACCA	660
TCCCAGTAAA GAACTTTAAA GCGGTCTTTA CGTCCACCAC AAAAGAGAAA GACTTGATCG	720
GAGAAAGGAT CCAATTCAAA GTGGGTTTGG	750

(2) INFORMATION FOR SEQ ID NO: 383:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 738 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 383:

TCAAATTCTT CGTGGTCCGC ATATCTnTCT TCGTACACGG CAGTCACTTG GTCTTCACT	60
ACTCGAGTCG CAGCTTCACG GGCCAATTTT TCTTCTACTT GAACTGCCTT TTGGAGGTCA	120
CTGTTGTAGG CTGCAATGAT TTCAGCTTGC AATTCAGCAT CCACGTGAAG CAATTCCACT	180
TCTGCTTTTT CTTTACCGAC AGCAGCAACG ATTTCTTCTT GGAAGGCAAT CAATTCTTTG	240
ACAGCTTCGT GCCCTTTAAG GAGCGCTTCC AACATGATTT CTTCTGACAA TTCTTTGGCA	300
CCAGACTCTA CCATGTTGAT AGCGTGCTTG GTTCCAGCTA CTGTCAATTC AAGAAGAGAT	360
TGCTCTGCTT GTTCTTGACT TGGGTTGATG ATGATTGGC CATCTACATA TCCCACTTGT	420
ACCCAGCAA TTGGTCCGTC AAATGGAATA TCTGAAATAG ACAGTGCCAA AGATGAACCA	480
AACATAGCAG CCATTGGTGC AGATGCATTT TCATCATAAG AAAGCACTGT ATTGATGACT	540
TGGACTTCAT TACGGAAACC TTCCGCAAAC ATAGGACGAA TCGGACGGTC AATCAAACGC	600
GCTGTCAAGG TCGCATCTGT TGAAGGACGT CCTTCACGTT TCATAAAGCC ACCAGGAAAC	660
TTCCCAGCCG CATACATTTT TTCTTCGTAG TTGACTTGA GTGGGAAGAA ATCCTCAGTT	720
GCCATTTTCT GGGGATCC	738

(2) INFORMATION FOR SEQ ID NO: 384:

1394

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 657 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 384:

CCCCCTATTT ACCGTGGACT AAAGTTGTAC AAGAAAAGTG CAAATAAGAA ATCTCCAGAT	60
TAGGAACATAT ATATGAGTTC TCTAGTCTGG AGATTTTCA ATAGACTTCG TTATTGGGCG	120
GTTACTTTTCG AAACTTTGAA AACTTCAAAA AACGGATTTT TATCGCTTTC AAATTCTTTT	180
GGGGTCAAAC TCAGTAACTT ATTCGCCTTG TAGACTTCAT GACGCTCAGG GTATACTTTC	240
AAGGTCCCAA ATAGCCAAGA ATCGTCAGCG ATATTATCTG AATCATCTCC TTCTTGTTCT	300
CCTTTAGTTC GCCTGAGGAC AGCCTTGACA CGCGCCAGAA TTCTCTAGGG CTAAAAGGCT	360
TGGTCAGGTA GTCATCAGCC CCTAATTCCA AGGCCAAAAC CTTATCAAAT TCATCACTTT	420
TCGCAGAAAC CATCATAATT GGAGTTTGA CGCCTTTGGC TCTCAGCCGC TTACAAACTT	480
CCATGCCATC TAATTGTGGT AACATGATAT CAAGCAAGAT AAAATCAAAG GGTTCGTGTT	540
CTGCCAAAGC TAAGGCCTTC CGTCCATTG TCACCAATTG AGTAGAAAAG CCTTCCTTAC	600
TTAAATGGTA GTCAAGCAAT TTCAGAATGT GTTCTTCATC ATCCACTAAT AAGACTT	657

(2) INFORMATION FOR SEQ ID NO: 385:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 586 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 385:

CCGCATCAGC ATCAACGAGT GCATCGGCTT CACGTCAACC AGTGCATCAG TCTCAGCAAG	60
CACCAGTGCG TCGGCTTCAG CATCAACGAG TGCCTCAGCC TCAGCAAGTA TCTCAGCGTC	120
TGAATCGGCA TCAACGAGTG CGTCAGCTCA GCAAGTACTA GTGCATCGGC TTCAGCAAGC	180
ACCAGTGCCT CGGCTTCAGC ATCAACGAGT GCCTCAGCCT CAGCAAGTAT CTCAGCGTCT	240
GAATCGGCAT CAACGAGTGC GTCACCTCAG CAAGTACTAG TGCATCAGCA TCAGCATCAA	300
CGAGTGCATC GGCTTCAGCA AGTACCAGCG CCTCAGCTTC AGCAAGCACC AGTGCCTCAC	360
CTCAGCAAGT ACCAGCGCCT CAGCCTCAGC AAGCACCAGT GCCTCAGCTT CAGCAAGTAC	420
CAGTGCCTCA CCTCAGCATC GACAAGTGGC TCGGCTTCAG CAAGTACCTC AGCGTCTGAA	480

1395

TCAGCATCAA CGAGTGCCTC AGCTTCAGCA TCAACCAAGTG CCTCAGCCTC AGCAAGTATC 540
AGTGCCTCAG CTTCAGCATC AACGAGTGCG TCAGCTGCAG CAAGTA 586

(2) INFORMATION FOR SEQ ID NO: 386:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 451 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 386:

CGTCGGCTTC AGCATCAACG AGTGCATCAG CTTCAGCATC AACAAGTGCT TCAGCTTCAG 60
CAAGTACCAG TGCCTCGGCT TCAGCATCAA CGAGTGCTTC AGTCTCAGCG TCAACCAAGTG 120
CCTCTGAATC CGCATCAACA AGTGCCTCGG CTTCAGCAAG CACCAGTGCT TCGGCTTCAG 180
CGTCAACGAG TGCCTCTGAG TCAGCATCAA CGAGTGCCTC ACCTCAGCAA GCACATCAGC 240
TTCTGAATCT GCATCAACCA GTGCGTCAGC TTCCGCATCA ACAAGCGCCT CGGCCTCAGC 300
AAGTACAAGT GCTTCAGCCT CAGCATCAAC CAGTGCATCA GCTTCAGCCT CAACAAGTGC 360
TTCAGCCTCA GCGTCAACCA GTGCCTCGGC TTCAGCAAGT ACCAGTGCCT CAGTTcAGCA 420
AGCACAAGTG CGTCAATTTA GCATCAACCA G 451

(2) INFORMATION FOR SEQ ID NO: 387:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 425 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 387:

TCTCAGCAAG CACCATTGCG TCGGCTTCAT CAAGCACCAG CGCGTTTGAA TCCGCATCAA 60
CCAGTGCTTC AGCTTCAGCC AAGTTACCTC AGCATCTGAA TCAGCATCAA CAAGTGCATC 120
GGCTTCAGCA AGCACAAGTG CTTCAGCTCA GCAAGTATCT CAGCGTCTGA ATCGGCATCA 180
ACGAGTGCCT CCGCTTCAGC AAGTACTAGC GCCTCAGCAT CAGCGTCAAC AAGTGTCTCG 240
GCTTCAGCGT CAACGAGTGC GTCTGAGTCA GCATCAACGA GTACGTCAGC CTCAGCAAGC 300
ACATCAGCTT CTGAATCTGC ATCAACCAAGT GCGTCAGCCT CAGCATCGAC AAGCGCCTCA 360
GCTTCAGCAA GTACCAAGTGC GTCAGCCTCA GCAAGTACCA GTGCTTCAGC CTCAGCGTCG 420

ACAAG 1396 425

(2) INFORMATION FOR SEQ ID NO: 388:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 572 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 388:

AGAGGATCCC CGGATCCTCA GTCGCTGAGA TAACTCCTTT GGGCTTGTTT ATCATGTAGT 60
AGACAAACTC TTCATACTCC AACACTTGCC CATTTTATGC GAATCTCATC TATTTTCTCT 120
TTTTTTTGCA ATTTAGCTGA TTTTCTTTT TTACCATTTA CAGTCACGCG CCCAGCCTTG 180
AGCAAGTTTT TGACCTCAGT CCGACTTCCC ACCGCACAGG CAACTAAAAA TTTATCTAAT 240
CTCATAGAAC TATTATATCA TATCAAAAGG AGGCTAGTAC AATGACCAAC CTCCTTTTCG 300
TTTCATACTC TTCAAAAATC TCTTCAAACC GCGTCAACGT CGCCTTGCCG TATATATGTT 360
ACTGACTTCG TCAGTTCTAT CTGCAACCTC AAAACAGTGT TTTGAGCTGA CTTCGTCAGT 420
TCTATCTGCA ACCTCAAAGC AGTGCTTTGA GCATCCTGCG GCTAGTTTCC kAGTkTGCTC 480
TTTGATTTC ATTGAGTATC AGATTTAGGA AATTAAGTTC CTCGKCTCCA AAAAAKAGCT 540
AAAACAATCA AGGCTCCTAA AATCGCTGGG AT 572

(2) INFORMATION FOR SEQ ID NO: 389:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 505 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 389:

CAACAAGTGC CTCGGCTTCA GCATGCACAA GTGCTTCAGC TTCAGCATGT ACCTGAGCGT 60
CTGAATCAGC ATCAACGTGT GCGTCCGCTT CAGCATGTAC TGCTGCCTCA GCATCAGCGT 120
CAACAAGTGC TTCGGCTTCA GCGTCAACGA GTGCGTCTGA GTCAGCATCA ACGAGTACGT 180
CAGCCTCAGC AAGCACATCA GCTTCTGAAT CTGCATCAAC CAGTGCGTCA GCCTCAGCAT 240
CGACAAGCGC CTCAGCTTCA GCAAGTACCA GTGCGTCAGC CTCAGCAAGT ACCAGTGCTT 300
CAGCCTCAGC GTCGACAAGT GCGTCGCGCT CAACCAAGTGC ATCTGAATCG GCATCAACCA 360
GTGCGTCAGC CTCAGCAAGT ACTAGCGCCT CAGCCTCAGC ATCAACGAGT GCGTCCGCTT 420

1397

CAGCAAGTAC TAGTGCATCA GCATCAGCAT CAACGAGTGC ATCGGCTTCA GCAAGTACCA 480
GCGCCTCAGC TTCAGCAAGC ACCGG 505

(2) INFORMATION FOR SEQ ID NO: 390:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 447 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 390:

GCTAAGACTA CCTCATTAGG GGCATAGGCT GCTAAAATAA CTGCAGCTGT GGTTAATGAC 60
AATACTGTAC TTTTTCAT TTTAATTCCT TACATATTTA TATAACTTCC AATAGATAAT 120
AACTTTAAC TTGCTAGCC TTGTATATAA AAAGTTTAC TAAGTATTAT CTAGGAAATA 180
GAGTAGTACA TTTATATATA ATTGTTATCT CTCTATAAAA ACAGTATATC ATTTAAAAA 240
ATTTAAGTCA AAAAAATTA CATTAGTTAA TTTATTTTT AGCACACATT AAAAAATAAG 300
ATTAGTACTC AATGAAAATC AAAGAGCAA CTAGGAAACT AGCCGCAGAT TGCTCAAAAC 360
AGTGTTTTGA GGTGTAGAT GGAATGACGT AGTCAGCTCA AAACACTGTT TTGAAGTTGT 420
GGATAGAACT GACGAAGTCG GTACCGA 447

(2) INFORMATION FOR SEQ ID NO: 391:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 572 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 391:

AGCACTTGTC GTTGAATTCT ACAACAAAT GTTGAATAT TTTATTGAAT AAGATAGGCC 60
TTGATATTAA GCACTTTGGG ACGTTCTCCC TTAGTGCTTT TTTGATTCT CTAGTATCC 120
AGCTATAATC GTTGAGACAT AACTAGACCG ATATAGTCCA AAGTGATATA GTAAAATGAA 180
CCAAAAATAG TACACAATGT GGTATAATCC TTTATGGCA TATTCAATAG ATTTTCGTAA 240
AAAAGTTCTC TCTTATTGTG AGCGAACAGG TAGTATAACA GAAGCATCAC ACGTTTCCCA 300
AATCTCACGT AATACCATT ATGGCTGGTT AAAGCTAAAA GAGAAAACAG GAGAGCTAAA 360
CCACCAAGTA TAGTGATTG AATCTATAAC AGTACACCTT GGCTGCTAAA ATATTCTAT 420

1398
AAATTAATTT GACTTTCCTG ATAGAGATGT TCACATCTTA TTCAAACCTA CTATATAAGT 480
TCTATAATCT CTTTATAAGA TTGCCCCATC AGACAAAATA GAACGATTG AAGGCGTTTA 540
TGATATTTAG CTGTACGAGA GTCTTTTAAA AG 572

MISSING UPON TIME OF PUBLICATION

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person approved by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PUT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the International publication of the application.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapse, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever two dates occurs earlier.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegians Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

ICELAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected in the art.

What Is Claimed Is:

- 25 1. Computer readable medium having recorded thereon the nucleotide sequence depicted in SEQ ID NOS:1-391, a representative fragment thereof or a nucleotide sequence at least 95% identical to a nucleotide sequence depicted in SEQ ID NOS:1-391.
- 30 2. Computer readable medium having recorded thereon any one of the fragments of SEQ ID NOS:1-391 depicted in Tables 2 and 3 or a degenerate variant thereof.
- 35 3. The computer readable medium of claim 1, wherein said medium is selected from the group consisting of a floppy disc, a hard disc, random access memory (RAM), read only memory (ROM), and CD-ROM.
- 40 4. The computer readable medium of claim 3, wherein said medium is selected from the group consisting of a floppy disc, a hard disc, random access memory (RAM), read only memory (ROM), and CD-ROM.
5. A computer-based system for identifying fragments of the *Streptococcus pneumoniae* genome of commercial importance comprising the following elements:
- 45 a) a data storage means comprising the nucleotide sequence of SEQ ID NOS:1-391, a representative fragment thereof, or a nucleotide sequence at least 95% identical to a nucleotide sequence of SEQ ID NOS:1-391;
- b) search means for comparing a target sequence to the nucleotide sequence of the data storage means of step (a) to identify homologous sequence(s), and
- c) retrieval means for obtaining said homologous sequence(s) of step (b).
- 50 6. A method for identifying commercially important nucleic acid fragments of the *Streptococcus pneumoniae* genome comprising the step of comparing a database comprising the nucleotide sequences depicted in SEQ ID NOS:1-391, a representative fragment thereof, or a nucleotide sequence at least 95% identical to a nucleotide sequence of SEQ ID NOS:1-391 with a target sequence to obtain a nucleic acid molecule comprised of a complementary nucleotide sequence to said target sequence, wherein said target sequence is not randomly selected.
- 55

60 7. A method for identifying an expression modulating fragment of
Streptococcus pneumoniae genome comprising the step of comparing a database
comprising the nucleotide sequences depicted in SEQ ID NOS:1-391, a
representative fragment thereof, or a nucleotide sequence at least 95% identical to
the nucleotide sequence of SEQ ID NOS:1-391 with a target sequence to obtain a
nucleic acid molecule comprised of a complementary nucleotide sequence to said
65 target sequence, wherein said target sequence comprises sequences known to
regulate gene expression.

70 8. An isolated protein-encoding nucleic acid fragment of the *Streptococcus*
pneumoniae genome, wherein said fragment consists of the nucleotide sequence of
any one of the fragments of SEQ ID NOS:1-391 depicted in Tables 2 and 3, or a
degenerate variant thereof.

75 9. A vector comprising any one of the fragments of the *Streptococcus*
pneumoniae genome SEQ ID NOS:1-391 depicted in Tables 2 and 3 or a
degenerate variant thereof.

80 10. An isolated fragment of the *Streptococcus pneumoniae* genome,
wherein said fragment modulates the expression of an operably linked open reading
frame, wherein said fragment consists of the nucleotide sequence from about 10 to
200 bases in length which is 5' to any one of the open reading frames depicted in
Tables 2 and 3 or a degenerate variant thereof.

85 11. A vector comprising any one of the fragments of the *Streptococcus*
pneumoniae genome of claim 8.

12. An organism which has been altered to contain any one of the
fragments of the *Streptococcus pneumoniae* genome of claim 8.

90 13. An organism which has been altered to contain any one of the
fragments of the *Streptococcus pneumoniae* genome of claim 10.

14. A method for regulating the expression of a nucleic acid molecule comprising the step of covalently attaching to said nucleic acid molecule a nucleic acid molecule consisting of the nucleotide sequence from about 10 to 100 bases 5' to any one of the fragments of the *Streptococcus pneumoniae* genome depicted in SEQ ID NOS:1-391 and Tables 2 and 3 or a degenerate variant thereof.

15. An isolated nucleic acid molecule encoding a homolog of any of the fragments of the *Streptococcus pneumoniae* genome of SEQ ID NOS:1-391 and Tables 2 and 3, wherein said nucleic acid molecule is produced by a process comprising steps of:

- a) screening a genomic DNA library using as a probe a target sequence defined by any of SEQ ID NOS:1-391 and Tables 2 and 3, including fragments thereof;
- b) identifying members of said library which contain sequences that hybridize to said target sequence; and
- c) isolating the nucleic acid molecules from said members identified in step (b).

16. An isolated DNA molecule encoding a homolog of any one of the fragments of the *Streptococcus pneumoniae* genome of SEQ ID NOS:1-391 and Tables 2 and 3, wherein said nucleic acid molecule is produced a process comprising steps of:

- a) isolating mRNA, DNA, or cDNA produced from an organism;
- b) amplifying nucleic acid molecules whose nucleotide sequence is homologous to amplification primers derived from said fragment of said *Streptococcus pneumoniae* genome to prime said amplification;
- c) isolating said amplified sequences produced in step (b).

17. An isolated polypeptide encoded by any of the fragments of the *Streptococcus pneumoniae* genome of SEQ ID NOS:1-391 and depicted in Table 2 and 3 or by a degenerate variant of said fragments.

18. An isolated polynucleotide molecule encoding any one of the polypeptides of claim 17.

19. An antibody which selectively binds to any one of the polypeptides of claim 17.

130

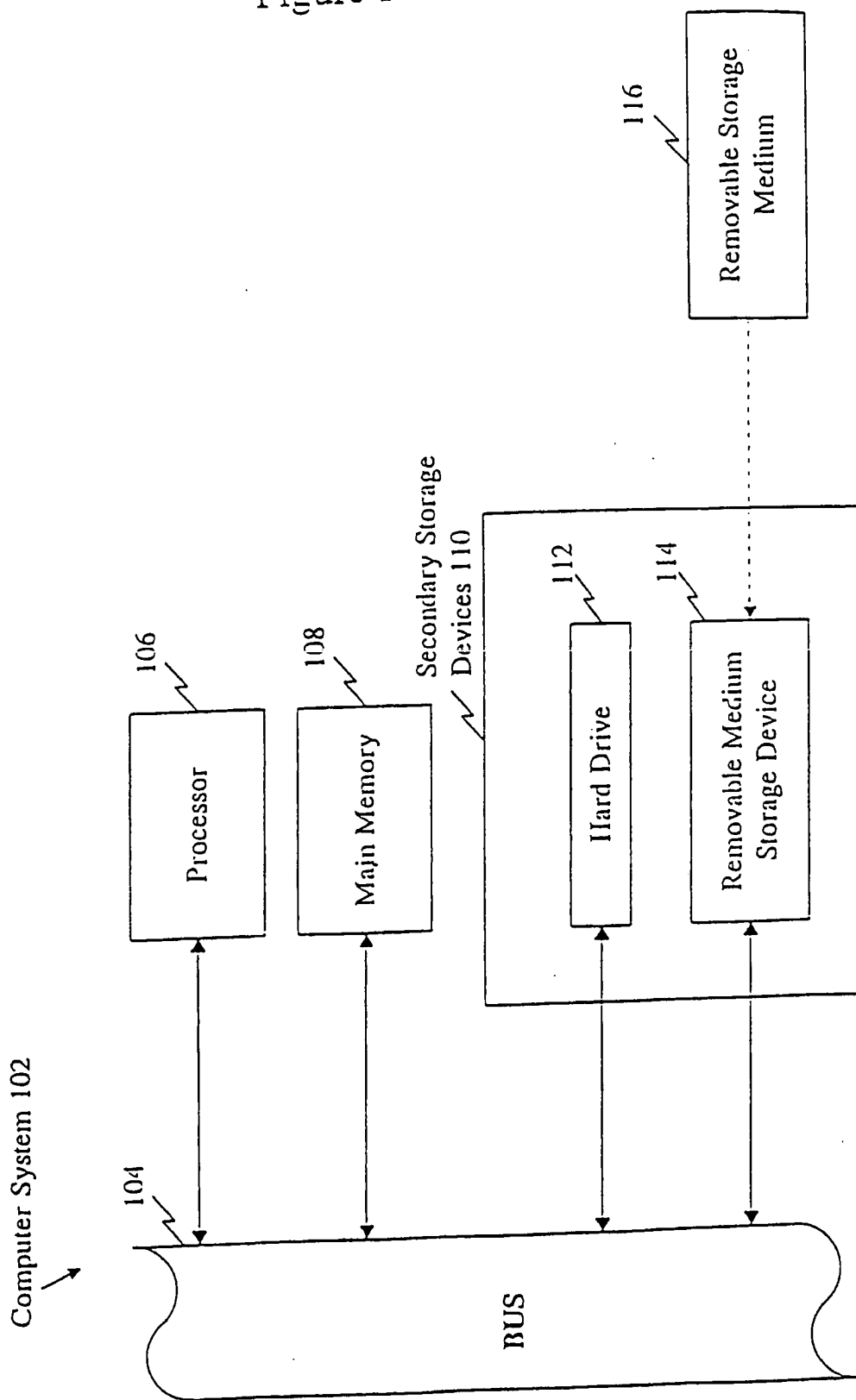
20. A method for producing a polypeptide in a host cell comprising the steps of:

135

a) incubating a host containing a heterologous nucleic acid molecule whose nucleotide sequence consists of any one of the fragments of the *Streptococcus pneumoniae* genome of SEQ ID NOS:1-391 and depicted in Tables 2 and 3, under conditions where said heterologous nucleic acid molecule is expressed to produce said protein, and

b) isolating said protein.

Figure 1



1384
AGTGCTTCAG TCTCAGCGTC AACCAGTGCC TCTGAATCAG CATCAACAAG TGCCTCGGCT 900
TCAGCAAGCA CCAGTGCGTC GGCTTCAGCA AGTACTAGTG C 941

(2) INFORMATION FOR SEQ ID NO: 369:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 869 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 369:

CAGCAAGTAC TAGTGACATCA GCTTCAGCAT CAACGAGTGC ATCGGCTTCT GCGTCAACCA 60
GTGCATCAGA GTCAGCAAGT ACCAGTGCGT CAGCTTCCGC ATCAACAAGT GCCTCGGCTT 120
CAGCAAGCAC CAGTGCGTCG GCTTCAGCAA GTACTAGCGC CTCAGCCTCA GCCTCAACCA 180
GTGCGTCAGC CTCAGCAAGT ATCTCAGCGT CTGAATCGGC ATCAACGAGT GCGTCCGCTT 240
CAGCAAGTAC TAGCGCCTCA GCCTCAGCGT CAACAAGTGC ATCGGCTTCA GCGTCAACGA 300
GTGCGTCTGA ATCGGCATCA ACGAGTGCGT CCGCTTCAGC AAGTACTAGC GCCTCAGCCT 360
CAGCGTCAAC AAGTGACATCG GCTTCAGCAT CAACGAGTGC GTCCGCTTCA GCAAGTACTA 420
GCGCTTCAGC CTCAGCGTCA ACAAGTGCGT CCGCTTCAGC GTCAACGAGT GCGTCTGAGT 480
CAGCATCAAC GAGTGCGTCA GCCTCAGCAA GCACATCAGC TTCTGAATCT GCATCAACCA 540
GTGCGTCAGC CTCAGCATCG ACAAGCGCCT CAGCTTCAGC AAGTACCAGT GCGTCAGCCT 600
CAGCGTCGAC AAGTGCGTCG GCTTCAGCAA GTACCAAGTGC GTCAGCCTCA GCAAGTACCA 660
GTGCGTCAGC CTCAGCGTCG ACAAGTGCGT CCGCCTCAAC CAGTGCACTT GAATCGGCAT 720
CAACCAAGTGC GTCAGCCTCA GCAAGTACTA GTGCATCAGC TTCAGCATCA ACGAGTGCGT 780
CGGCTTCAGC ATCAACCAAGT GCATCAGAGT CAGCAAGTAC CAGTGCGTCA GntTCCGCAT 840
GCAACAAGTG CCTCGGCTTC AGCAAGTAC 869

(2) INFORMATION FOR SEQ ID NO: 370:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 750 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 370:

TCAACAAGTG CCTCAGCATC AGCATCAACG AGTGCGTCAG CCTCAGCAAG TACTAGTGCA 60

1385

TCAGCATCAG CATCAACCAG TGCATCAGCC TCAGCAAGTA TCTCAGCGTC TGAATCGGCA	120
TCAACGAGTG CATCAGCATC AGCATCAACG AGTGCATCGG CTTCAGCGTC AACCAGTGCA	180
TCAGTCTCAG CAAGCACCAG TCGTCGGCT TCAGCATCAA CGAGTGCCTC AGCCTCAGCA	240
AGTATCTCAG CGTCTGAATC GGCATCAACG AGTGCCTCAG CCTCAGCAAG TACTAGTGCA	300
TCGGCTTCAG CAAGCACCAG TCGTCGGCT TCAGCATCAA CCAGTGCCTC AGCCTCAGCA	360
AGTATCTCAG CGTCTGAATC GGCATCAACG AGTGCCTCAG CCTCAGCAAG TACTAGTGCA	420
TCAGCATCAG CATCAACGAG TGCATCGGCT TCAGCAAGTA CCAGCGCCTC AGCTTCAGCA	480
AGCACCAGTG CGTCAGCCTC AGCAAGTACC AGCGCCTCAG CCTCAGCAAG CACCAGTGCC	540
TCAGCTTCAG CAAGTACCAG TCGTCAGCC TCAGCGTCGA CAAGTGCCTC GGCTTCAGCA	600
AGTACCTCAG CGTCTGAATC AGCATCAACG AGTGCATCAG CTTCAGCATC AACAAAGTGCT	660
TCAGCTTCAG CAAGTATCTC AGCGTCTGAA TCGGCATCAA CGAGTGCCTC CGCTTCAGCA	720
AGTACTAGCG CCTCAGCATC AGCGTCAACG	750

(2) INFORMATION FOR SEQ ID NO: 371:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 957 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 371:

CCGGAAAACA GCTCTGGCGC TTGGTCTTGC CCAGCGTATT GCTAGTGGTG ACGTGCCTGC	60
GGAAATGGCT AAGATGCGCG TGTTAGAACT TGATTTGATG AATGTCGTG CAGGGACACG	120
CTTCCGTGGT GACTTTGAAG AACGCATGAA TAATATCATC AAGGATATTG AAGAAGATGG	180
CCAAGTCATC CTCTTTATCG ATGAACTCCA CACCATCATG GGTTCCTGTA GCGGGATTGA	240
TTCGACTCTG GATGCGGCCA ATATCTTGAA ACCAGCCTTG GCGCGTGGAA CTTTGAGAAC	300
GGTTGGTGCC ACTACTCAGG AAGAATATCA AAAACATATC GAAAAAGATG CGGCACTTTC	360
TCGTCGTTTC GCTAAAGTGA CGATTGAAGA ACCAAGTGTG GCAGATAGTA TGA TATTTT	420
ACAAGGTTTG AAGGCGACTT ATGAGAAACA TCACCGTGT CAAATCACAG ATGAAGCGGT	480
TGAAACAGCG GTTAAGATGG CTCATCGTTA TTAAACAGT CGTCACTTGC CAGACTCTGC	540
TATCGATCTC TTGGATGAGG CGGCAGCAAC AGTGCAAAT AAGGCAAAGC ATGTAAAAGC	600
AGACGATTCA GATTTGAGTC CAGCTGACAA GGCCCTGATG GATGGCAAGT GGAAACAGGC	660

1386

AGCCCAGCTA ATCGCAAAAG AAGAGGAAGT ACCTGTCTAC AAAGACTTGG TGACAGAGTC	720
TGATATTTTG ACCACCTTGA GTCGCTTGTC AGGAATCCCA GTTCAAAAAC TGA CTCAAAC	780
GGATGCTAAG AAGTATTTAA ATCTTGAAGC AGAACTCCAT AAACGGGTTA TCGGTCAAGA	840
TCAAGCTGTT TCAAGCATTG GCCGTGCCAT TCGCCGCAAC CAGTCAGGGA TTCGCAGTCA	900
TAAGCGTCCG ATTGGTTCCT TTATGTTCCT AGGGCCTACA GGTGTCGGGG TATCCGA	957

(2) INFORMATION FOR SEQ ID NO: 372:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 807 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 372:

CAAAGCGCCT CAGCTTCAGC ATCAACAAGT GCGTCGGCTT CAGCATCAAC CAGTGCCTCG	60
GCTTCAGCGT CAACCAAGTGC GTCACATTCA GCAAGTACCA GTGCTTCAGT CTCAGCATCA	120
ACAAGTGCTT CAGCCTCAGC ATCGACAAGT GCCTCGGCTT CAGCAAGCAC ATCAGCATCT	180
GAATCAGCGT CAACCAAGTGC TTCGGCTTCA GCAAGTACCA GTGCTTCAGC TTCAGCATCA	240
ACCAGCGCCT CGGCCTCAGC AAGCACCTCA GCTTCTGAAT CGGCCTCAAC CAGCGCCTCG	300
GCCTCAGCAA GCACCTCAGC TTCTGAATCG GCCTCAACCA GCGCCTCAGC CTCAGCATCA	360
ACGAGTGCTT CGGCTTCAGC AAGCACAAAG GCCTCGGGTT CAGCATCAAC GAGTACGTCA	420
GCTTCAGCGT CAACCAAGTGC TTCAGCCTCA GCATCAACAA GTGCGTCAGC CTCAGCAAGT	480
ATCTCAGCGT CTGAATCGGC ATCAACGAGT GCGTCTGAGT CAGCATCAAC GAGTACGTCA	540
GCCTCAGCAA GCACCTCAGC TTCTGAATCG GCCTCAACCA GTGCGTCAGC CTCAGCATCG	600
ACAAGCGCCT CAGCTTCAGC AAGTACCAAG GCTTCAGCCT CAGCGTCGAC AAGTGCCTCG	660
GCCTCAACCA GTGCATCTGA ATCGGCATCA ACCAGTGCCT CAGCCTCAGC AAGTACTAGT	720
GCATCGGCTT CAGCATCAAC CAGTGCCTCG GCTTCAGCGT CAACCAAGTGC GTCAGCTTCA	780
GCAAGTACCA TGTGCTTCAT GTCTCAG	807

(2) INFORMATION FOR SEQ ID NO: 373:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1068 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

1387

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 373:

CATCGGCTTC AGCATCAACG AGTGCCTCCG CTTAGCAAG TACTACCGCC TCAGCCTCAG	60
CGTCAACAAG TGCATCGGCT TCAGCGTCAA CGAGTGCCTC TGAGTCAGCA TCAACGAGTG	120
CGTCACCTCA GCAAGCACAT CAGCTTCTGA ATCTGCATCA ACCAGTGCCT CACCTCAGCA	180
TCGACAAGCG CCTCAGCTTC AGCAAGTACC AGTGCCTCAC CTCAGCGTCG ACAAGTGCCT	240
CGGCTTCAGC AAGTACCAGT GCGTCAsCTC AGCAAGTACC AGTGCCTCAC CTCAGCGTCG	300
ACAAGTGCCT CGGCCTCAAC CAGTGCATCT GAATCGGCAT CAACCAAGTGC GTCACCTCAG	360
CAAGTACTAG TGCATCAGCT TCAGCATCAA CGAGTGCATC GGCTTCAGCA TCAACCAAGTG	420
CATCAGAGTC AGCAAGTACC AGTGCCTCAG cTTCCGCATC AACCAAGTGCC TCGGCTTCAG	480
CAAGTACTAG CGCCTCAGCC TCAGCGTCAA CAAGTGCTTC AGCTTCCGCG TCAACCAAGCG	540
CCTCGGCCTC AGCAAGTATC TCAGCGTCTG AATCGGCATC AACCAAGTGCC TCGGCTTCAG	600
CATCAACGAG TGCATCAGTC TCAGCAAGCA CCAAGTGCCTC GGCCTCAGCA AGCACCAGCG	660
CGTCTGAATC CGCATCAACC AGTGCCTCAG CTTAGCAAG TACTAGCGCC TCAGCATCAG	720
CATCAACAAG TGCATCGGCT TCAGCAAGCA CAAGTGCTTC AGCCTCAGCA AGTATCTCAG	780
CGTCTGAATC GGCATCAACG AGTGCCTCCG CTTAGCAAG TACTAGCGCC TCAGCATCAG	840
CGTCAACAAG TGCTTCGGCT TCAGCGTCAA CGAGTGCCTC TGAGTCAGCA TCAACGAGTA	900
CGTCAGCCTC AGCAAGCACA TCAGCTTCTG AATCTGCATC AACCAGTGCCT TCAGCCTCAG	960
CATCGACAAG CGCCTCAGCT TCAGCAAGTA CCAAGTGCCTC AGCCTCAGCA AGTACCAGTG	1020
CTTCAGCCTC AGCGTCGACA AGTGCCTCGG GCTCAACCAG TGCACTCTG	1068

(2) INFORMATION FOR SEQ ID NO: 374:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 620 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 374:

CAGCATCAAC GAGTGCTTCA GTTTCAGCGT CAACCAAGTGC CTCTGAATCA GCTTCAACAA	60
GTGCCTCGGC TTCAGCAAGC CCCAGTGCCT CGGCTTCAGC AAGTACTAGT GCATCGGCTT	120
CAGCATCGAC AAGTGCCTCT GAATCGGCAT CAACGAGTGC TTCGGCTTCA GCATCAACGA	180
GTGCGTCAGC CTCAGCAAGC ACATCAGCTT CTGAATCTGC ATCAACCAGT GCGTCCGyTT	240

1388

CAGCGTCAAC CAGTGCCTCG GCTTCAGCGT CGACAAGTGC TTCGGCTTCA GCATCAACGA	300
GTGCGTCGGC CTCAGCAAGC GCAAGTACCT CAGCGTCAGC TTCCGCCTCA ACCAGTGCGT	360
CGGCTTCAGC AAGCACAACT GCGTCAGCCT CAGCAAGTAT CTCAGCGTCT GAATCGGCAT	420
CAACGAGTGC GTCTGAGTCA GCATCAACGA GTACGTCAGC CTCAGCAAGC ACATCAGCTT	480
CTGAATCTGC ATCAACCACT GCGTCAGCCT CAGCATCGAC AAGCGCCTCA GCTTCAGCAA	540
GTACCACTGC TTCAGCCTCA GCGTCGACAA GTGCGTCGGC CTCAACCACT GCATCTGAAT	600
CGGCATCAAC CAGTGCCTCA	620

(2) INFORMATION FOR SEQ ID NO: 375:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 720 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 375:

GTATTGGGGC GCCCCAACCT CTATGTGACT ACGGATTATT TCCTAGATTA CATGgGGATA	60
AACCATTTAG AAGAATTACC AGTGATTGAT GAGCTTGAGA TTCAAGCCCA AGAAAGCCAA	120
TTATTTGGTG AAAGGATAGA AGAAGATGAG AATCAATAAG TATATTGCCC ACGCAGGTGT	180
GGCCAGTAGG AGAAAAGCAG AAGAGCTGAT TAAGCAAGGC TTGGTGACGG TTAACGGCCA	240
AGTGGTGCGT GAACTAGCAA CCACTATCAA GTCAGGCGAC AAGGTCGAAG TTGAAGGTCA	300
ACCTATCTAC AACGAAGAAA AGGTCTACTA TCTGCTTAAC AAACCACGCG GTGTGATTTC	360
CAGTGTGACA GATGATAAGG GTCGCAAGAC GGTGTGTCGAC CTCTTGCCCA ATGTCAAAGA	420
GCGTATTTAC CCTGTGGGTC GTTTGGACTG GGATACATCA GGTGTCTTGA TTTTGACCAA	480
TGATGGGGAC TTTACAGACG AGATGATTCA CCCTCGTAAT GAGATTGACA AGGTTTATGT	540
CGCGCGTGT AAAGGTGTGG CCAATAAGGA CAATCTCCGC CCCTTGACCC GTGGTCTTGA	600
GATTGATGGT AAGAAAACCA AGCCATAATA TATAGGTTTT GTAGCCTCTA CACCATAAAT	660
ATTTGCTAAT AAAAATACTG TATTATTACC CTCTTAAGGT GCGAAATTAT TCAAGTTCTT	720

(2) INFORMATION FOR SEQ ID NO: 376:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 648 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

1389

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 376:

CGCCATTTCC CATCGTACCG CCGAAAATCC CAGCGCCTCA GCCATCAAAT ATCCTATCAA	60
CGTTCTCAAA AAAAGTGACC GCTCTCTCAT CATGTTTCCA AGTGGTAGCC GCCACTCAAA	120
CGATGTCAAG GGGGGCGCAC ACTskATTGC CAAAATGGCC AAGGTCCGTA TCATGCCGGT	180
TACCTACACC GGTCCCATGA CTTTGAAGGG CTTGATTAGC CGTGAACGTG TCGATATGAA	240
CTTTGAAAT CCAATCGATA TCTCAGATAT CAAGAAAATG AATGATGAAG GCATTGAAAC	300
AGTCGCCAAT CGTATTCAAA CAGAATTCCA ACGTCTGGAC GAAGAAACGA AACATGGCA	360
CAATGATAAA AAACCAAATC CACTCTGGTG GTTTATCCGC ATCCCTGCC TCATCCTTGC	420
TATTATCCTC GCTATCCTAA CCATCATCTT TAGCTTTATC GCAAGCTTCA TCTGGAACCC	480
AGATAAGAAA AGAGAAGAAC TTGCATAGAA GAAATGAACC TTGGCCAAAC AGCTAAGGTT	540
TTCATTTATA TAGTAGATTG GwACTAGAAT AGTACACCTC TACTTCTAAA ACATTTTATG	600
AAATCGATTT GACTGTCCTG ATCGATTTGT CCTAATCTTA TTTCAATT	648

(2) INFORMATION FOR SEQ ID NO: 377:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 690 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 377:

GTGCATCGCT TTCAGCATCG ACAAGTCCGT CTGAATCGGC ATCAACGAGT GCTTCGGCTT	60
CAGCATCAAC GAGTGCGTCA GCTTCAGCAA GCACATCAGC TTCTGAATCT GCATCAACCA	120
GTGCGTCCGC TTCAGCGTCA ACCAGTCCGT CGGCTTCAGC GTCGACAAGT GCTTCGGCTT	180
CAGCATCAAC GAGTGCGTCG GCCTCAGCAA GCGCAAGTAC CTCAGCGTCA GCTTCCGCCT	240
CAACCAAGTGC GTCCGCTTCA GCAAGCAGAA GTGCGTCAGC CTCAGCAAGT ATCTCAGCGT	300
CTGAATCGGC ATCAACGAGT GCGTCGGCCT CAGCAAGCGC AAGTACCTCA GCGTCAGCTT	360
CCGCCTCAAC CAGTGCGTCG GCTTCAGCAA GCACAAGTGC GTCAGCCTCA GCAAGTATCT	420
CAGCGTCTGA ATCGGCATCA ACGAGTCCGT CTGAGTCAGC ATCAACGAGT ACGTCAGCCT	480
CAGCAAGCAC ATCAGCTTCT GAATCGGCAT CAACCAAGTGC GTCAGCCTCA GCATCGACAA	540
GCGCCTCAGC TTCAGCAAGT ACCAGTGCTT CAGCCTCAGC GTCGACAAGT GCGTCGGCCT	600
CAACCAAGTGC ATCTGAATCG GCATCAACCA GTGCGTCAGC CTCAGCAAGT ACTAGTGCAT	660

CAGCTTCAGC ATCAACGAGT GCATCGGCTT 1390 690

(2) INFORMATION FOR SEQ ID NO: 378:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1003 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 378:

CGAGATTCTC TGGAGTTATG GATGTCGTTC CAATATGTGC ACGTTGGAAT GTTAGTGCTT 60
ATATGGGGGG AACAGAATCC TCTCTTGATT GAAGACAAGC TAGTCATTAG GCTGGTTTGT 120
CTTTTGTCA ACTGTAGTGG GTTGATATAA TAGTATTAGT GAGTGGGATA AAAGTTTCAT 180
TTAGTTTATT CAGTACAAAT TTAACGGGTC AAGATTTATA TACTAGTGGT GTTTTGGGG 240
CTGAGAGAAG TATCTTGATT TTATGTGTGG TTTTATACT TACAGTTGTT CTGCTCCAAA 300
GAGCTTGTAG AGAAGAATTA GCTCATAAAG GAGATTGATT ATTTTGATAT CAAAAAATG 360
CACAGATAA CCTGATGCAT TTTTTAGCG ACAATGCTTG CTACTTCCTT CTGTCGAATT 420
TAGACAATTT TAAACCCCAA TTATTCACCC CAAATCTAAA AACCATCCAG AATCCTTGCC 480
TTAGCTTAGA TCCTGGATGG TTTCTTTTTT CACCCAATGG GTGTTTTTTA CTAGACAAAA 540
AAGAGTTTCC CCTTTATGGT ATAAGTGTAG AAAAAACAC AAAAGAAAG GAAACTCACA 600
TGAACAGTTT ACCAAATCAT CACTTCCAAA ACAAGTCTTT TTACCAACTA TCTTTCGATG 660
GAGGTCATTT AACCCAGTAT GGTGGTCTTA TCTTTTTTCA GGAACTTTTT TCCCAGTTGA 720
AACTAAAAGA GCGGATTCTT AAGTATTTAG TAACGAATGA CCAACGCCGC TACTGTCGTT 780
ATTGCGGATTC AGATATCCTT GTCCAGTTCC TCTTCAACT GTTAACAGGT TATGGAACGG 840
ACTATGCTTG TAAAGAATTG TCAGCTGATG CCTACTTTCC AAAATTATTG GAAGGAGGGC 900
AGCTTGCTTC ACAGCCAACC TTATCCCGTT TTCTTCCAG AACTGACGAG GAAACAGTCC 960
ATAGTTTGGC ATGCCTCAAC CTTGAATTGG TCGAATTCTT TTT 1003

(2) INFORMATION FOR SEQ ID NO: 379:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 738 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 379:

1391

CCGATGATTC TGATTGGTTT GCTCTTTACT TTGCTGGGAA TTTTGAGGTA GATCTATGAT	60
TGAAATACTA ATTGTTTTAG CTATTATCCT ATCTCTTGCT TTGATTGTAT TGGTAACTAT	120
ACAACCCCGT CAAAATCAAC TATTTTCCAT GGATGCCACT AGTAATATTG GTAAACCAAG	180
CTACTGGCAG AGCAACACCT TGGTCAAGGT GCTCACTTTA TTGGTGAGTT TGGCTTTATT	240
TATTCTACTA TTAACCTTTA TGGTGATTAC TTATAAATAA AAGAAAACCT CAGATATTCA	300
CCTTTTGTGG ATTGGTCTGA AGTTTCTCTT TTTATACTCA ATGAAAATCA AAGAGCAAAC	360
TAGGAAGCTA GCCGCackGC TCAAAACACC GTTTTGAGGT TGTAGATATA ACTGACGAGc	420
GACTCAAAAC ACCGTTTGA GGTGTAGAT ATAAGTACG AGcGACTCAA AACACCGTTT	480
TGAGGTGTG GATAGAACTG ACGAGcGACT CAAACACCG TTTTGAGGTT GTGGATAGAA	540
CTGACGAAGT CGcTCAAAAC ACCGTTTGA GGTGTGGAT AGAACTGACG AatgctCAAA	600
ACACCGTTTT GAGGTGTGG ATAGAACTGA CGAAGCgaaC ATATATACAG CAAGGCGACG	660
CTGACGTGGT TTGAAGAGTA TTACTGTCTA TATTTTGGT AAAAATCAAC TTTTACTTGG	720
ATGAAGGTTT TTTTTTTT	738

(2) INFORMATION FOR SEQ ID NO: 380:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 695 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 380:

CCGTCTTATC AAAGAGGTTA ACAAAGGCAC CAAATTCTC GATACGAACG ACTTTAGCAC	60
GGTAAACTTC ATCCACTTTG GCTTCACGAA CCAAACCAGC AATAATTCTT TTGGCACGGT	120
TAATAGCATC TTGGTCACTA GAGTAGATAG ACACATTTC TTCTTCGTCT ATATCAATCT	180
TAACACCTGT TTCAGCGATA ATCTTGTCGA TGGTTTCTCC ACCCTTACCG ATGACAATCT	240
TAATCTTGTC CACATCAATC TTGATCGTAT CAATTTTCGG AGCAGTTGGA GCCAATTCTG	300
GACGAACTTC TGGAAATGGT GCTTCAATGA CATCAAGGAT TTCAAAACGC GCTTCTTGG	360
CTTGAGCAAG AGCCTCCGTC AAGATTCTG CAGTAATCCC TTGAATCTTG ATATCCATTT	420
GAAGGGCTGT AATCCCATCA CGAGTACCTG CAACCTTGAA GTCCATATCT CCAAAGTGAT	480
CTTCCAAACC TTGGATATCT GTCAATACTG TGTAGTTATT TCCATCTGAG ATAAGCCCCA	540
TAGCAATACC AGCTACTGGC GCCTTGATTG GCACACCACC AGCCATAAGG GCAAGAGTTC	600

1392

CCGCACAGAT AGAAGCTTGA GATGAAGAAC CGTTTGATTC CAAAACCTTCT GCTACTAGAC 660

GGATAGCGTA GGGGAATTCT TCCAAGCTTG GCAGG 695

(2) INFORMATION FOR SEQ ID NO: 381:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 691 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 381:

GACATCTTAT CTAAATACAT GCTAATATAT TTAGATACAA ACATTCCAAC TTGATAATTT 60

TCACTCATCT TTCATCATTC CTTATACAAC TATGCAGTAT AAATAGAATA GTTTTCTCAT 120

CAGAATGAGA CTATTTTAAT ATTAGATCCC CAATTATTCA CCCCAAATCT AAAAACCATC 180

CAGAATCCTT GCCTTAGCTT AGATCCTGGA TGGTTTCTTT TTTCACCCAA TGGGTGTTTT 240

TTACTAGACA AAAAAGAGTT TCCCCTTTAT GGTATAAGTG TAGAAAAAAA CACAAAAAGA 300

AAGGAACTC ACATGAACAG TTTACCAAAT CATCACTTCC AAAACAAGTC TTTTACCAA 360

CTATCTTTCG ATGGAGGTCA TTTAACCAG TATGGTGGTC TTATCTTTTT TCAGGAACTT 420

TTTTCCAGT TGAAACTAAA AGAGCGGATT TCTAAGTATT TAGTAACGAA TGACCAACGC 480

CGCTACTGTC GTTATTCGGA TTCAGATATC CTTGTCCAGT TCCTCTTTCA ACTGTTAACA 540

GGTTATGGAA CGGACTATGC TTGTAAAGAA TTGTCAGCTG ATGCCTACTT TCCAAAATG 600

TTGGAAGGAG GGCAGCTTGc TTCACAGCCA ACCTTATCCC GwTTTCTTTC CAGAACTGAC 660

GAGGAAACAG TCCATAGTTT GCGATGCCTC A 691

(2) INFORMATION FOR SEQ ID NO: 382:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 750 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 382:

ATCTCTCTGC GTAATGGTCC TCAGATAACT CTGATGATGT GTGGCGATAT AGAACTGAGC 60

CAAGTTATGC CTAAGGGGCC TTAGGAATAG GAGCTTTCAC AAGCTTATCC AGATGATTAT 120

CTTTTACTCG TTATGGACAA TGCTATATGG CATAAATCAA GTACCTTAAA GATTCCGACT 180

AATATTGGCT TTGCATTTAT TCCTCCATAC ACACCAGAGA TGAACCCCAT TGAACAAGTG 240

1393

TGGAAGAGA TTCGTAAACG TGGATTAAAG AATAAAGCCT TTCGAACTTT GGAAGATGTC	300
ATACAAGGAC TGGAGAAGGA GGTGATAAAG TCCATCGTTA ATCGGAGACG GACTAGAATG	360
CTTTTGTAAA ACAGATGAGT ATAAAAAGAA AGTCCTCATT TCAATAGAAA TCACGACTTT	420
CTGATGAATT TATAGTAAAA TGAAATAAGA ACAGGATAGT CAAATCGATT TCTAACAATG	480
TTTGTAGAAGC AGAGGTGTAC TATTCTAGTT TAAATCCACT ATATTTGGGG AGTGATAGAA	540
AAGCCCTTCA TCAGCCAATC TACTTGTTCA GGTGCGAGAG CTTTGACATC CTTTCTGTGA	600
CTGGACCAAG TCAGTTTTCG GTTCTCAAAG CGTTTATATA ATATCCAAAA TCCTTGACCA	660
TCCCAGTAAA GAACTTTAAA GCGGTCTTTA CGTCCACCAC AAAAGAGAAA GACTTGATCG	720
GAGAAAGGAT CCAATTCAAA GTGGGTTTGG	750

(2) INFORMATION FOR SEQ ID NO: 383:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 738 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 383:

TCAAATTCTT CGTGGTCCGC ATATCTnTCT TCGTACACGG CAGTCACTTG GTCTTTCCT	60
ACTCGAGTCG CAGCTTCACG GGCCAATTTC TCTTCTACTT GAACTGCCTT TTGGAGGTCA	120
CTGTGTAGG CTGCAATGAT TTCAGCTTGC AATTCAGCAT CCACGTGAAG CAATTCCACT	180
TCTGCTTTT CTTTACCGAC AGCAGCAACG ATTTCTTCTT GGAAGGCAAT CAATTCCTTG	240
ACAGCTTCGT GCCCTTTAAG GAGCGCTTCC AACATGATTT CTTCTGACAA TTCCTTGGA	300
CCAGACTCTA CCATGTTGAT AGCGTGCTTG GTTCCAGCTA CTGTCAATTC AAGAAGAGAT	360
TGCTCTGCTT GTTCTTGACT TGGGTTGATG ATGATTTGGC CATCTACATA TCCCCTTGT	420
ACCCAGCAA TTGGTCCGTC AAATGGAATA TCTGAAATAG ACAGTGCCAA AGATGAACCA	480
AACATAGCAG CCATTGGTGC AGATGCATTT TCATCATAAG AAAGCACTGT ATTGATGACT	540
TGGACTTCAT TACGGAAACC TTCCGCAAC ATAGGACGAA TCGGACGGTC AATCAAACGC	600
GCTGTCAAG TCGCATCTGT TGAAGGACGT CCTTCACGTT TCATAAAGCC ACCAGGAAAC	660
TTCCAGCCG CATACATTTT TTCTTCGTAG TTGACTTGGA GTGGGAAGAA ATCCTCAGTT	720
GCCATTTTCT GGGGATCC	738

(2) INFORMATION FOR SEQ ID NO: 384:

1394

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 657 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 384:

CCCCCTATTT ACCGTGGACT AAAGTTGTAC AAGAAAAGTG CAAATAAGAA ATCTCCAGAT	60
TAGGAACATAT ATATGAGTTC TCTAGTCTGG AGATTTTCA ATAGACTTCG TTATTGGGCG	120
GTTACTTTCG AAACTTTGAA AACTTCAAAA AACGGATTTT TATCGCTTTC AAATTCTTTT	180
GGGGTCAAAC TCAGTAACCTT ATTCGCCTTG TAGACTTCAT GACGCTCAGG GTATACTTTC	240
AAGGTCCCAA ATAGCCAAGA ATCGTCAGCG ATATTATCTG AATCATCTCC TTCTTGTTCT	300
CCTTTAGTTC GCCTGAGGAC AGCCTTGACA CGCGCCAGAA TTCTCTAGGG CTAAAAGGCT	360
TGGTCAGGTA GTCATCAGCC CCTAATTCCA AGGCCAAAAC CTTATCAAAT TCATCACTTT	420
TCGCAGAAAC CATCATAATT GGAGTTTGA CGCCTTTGGC TCTCAGCCGC TTACAAACTT	480
CCATGCCATC TAATTGTGGT AACATGATAT CAAGCAAGAT AAAATCAAAG GGTTCGTGTT	540
CTGCCAAAGC TAAGGCCTTC CGTCCATTG TCACCAATTG AGTAGAAAAG CCTTCCTTAC	600
TTAAATGGTA GTCAAGCAAT TTCAGAAATG GTTCTTCATC ATCCACTAAT AAGACTT	657

(2) INFORMATION FOR SEQ ID NO: 385:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 586 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 385:

CCGCATCAGC ATCAACGAGT GCATCGGCTT CACGTCAACC AGTGCATCAG TCTCAGCAAG	60
CACCAAGTGC TCGGCTTCAG CATCAACGAG TGCCTCAGCC TCAGCAAGTA TCTCAGCGTC	120
TGAATCGGCA TCAACGAGTG CGTCAGCTCA GCAAGTACTA GTGCATCGGC TTCAGCAAGC	180
ACCAGTGCCT CGGCTTCAGC ATCAACCACT GCCTCAGCCT CAGCAAGTAT CTCAGCGTCT	240
GAATCGGCAT CAACGAGTGC GTCACCTCAG CAAGTACTAG TGATCAGCA TCAGCATCAA	300
CGAGTGCATC GGCTTCAGCA AGTACCAGCG CCTCAGCTTC AGCAAGCACC AGTGCCTCAC	360
CTCAGCAAGT ACCAGCGCCT CAGCCTCAGC AAGCACCAGT GCCTCAGCTT CAGCAAGTAC	420
CAGTGCCTCA CCTCAGCATC GACAAGTGGC TCGGCTTCAG CAAGTACCTC AGCGTCTGAA	480

1395

TCAGCATCAA CGAGTGGCTC AGCTTCAGCA TCAACCAGTG CCTCAGCCTC AGCAAGTATC 540
AGTGGCTCAG CTTCAGCATC AACGAGTGGC TCAGCTGCAG CAAGTA 586

(2) INFORMATION FOR SEQ ID NO: 386:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 451 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 386:

CGTCGGCTTC AGCATCAACG AGTGCATCAG CTTCAGCATC AACAAAGTGCT TCAGCTTCAG 60
CAAGTACCAG TGGCTCGGCT TCAGCATCAA CGAGTGCTTC AGTCTCAGCG TCAACCAGTG 120
CCTCTGAATC CGCATCAACA AGTGCCTCGG CTTCAGCAAG CACCAGTGCT TCGGCTTCAG 180
CGTCAACGAG TGGCTCTGAG TCAGCATCAA CGAGTGGCTC ACCTCAGCAA GCACATCAGC 240
TTCTGAATCT GCATCAACCA GTGCGTCAGC TTCCGCATCA ACAAGCGCCT CGGCCTCAGC 300
AAGTACAAGT GCTTCAGCCT CAGCATCAAC CAGTGCATCA GCTTCAGCCT CAACAAGTGC 360
TTCAGCCTCA GCGTCAACCA GTGCCTCGGC TTCAGCAAGT ACCAGTGGCT CAGTTCAGCA 420
AGCACAAGTG CGTCAATTTA GCATCAACCA G 451

(2) INFORMATION FOR SEQ ID NO: 387:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 425 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 387:

TCTCAGCAAG CACCATTGCG TCGGCTTCAT CAAGCACCAG CGCGTTTGAA TCCGCATCAA 60
CCAGTGCTTC AGCTTCAGCC AAGTTACCTC AGCATCTGAA TCAGCATCAA CAAGTGCATC 120
GGCTTCAGCA AGCACAAGTG CTTCAGCTCA GCAAGTATCT CAGCGTCTGA ATCGGCATCA 180
ACGAGTGGCT CCGCTTCAGC AAGTACTAGC GCCTCAGCAT CAGCGTCAAC AAGTGCTTCG 240
GCTTCAGCGT CAACGAGTGC GTCTGAGTCA GCATCAACGA GTACGTCAGC CTCAGCAAGC 300
ACATCAGCTT CTGAATCTGC ATCAACCAGT GCGTCAGCCT CAGCATCGAC AAGCGCCTCA 360
GCTTCAGCAA GTACCAAGTGC GTCAGCCTCA GCAAGTACCA GTGCTTCAGC CTCAGCGTCG 420

ACAAG

1396

425

(2) INFORMATION FOR SEQ ID NO: 388:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 572 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 388:

AGAGGATCCC CGGATCCTCA GTCGCTGAGA TAACTCCTTT GGGCTTGTTT ATCATGTAGT	60
AGACAAACTC TTCATACTCC AACACTTGCC CATTTTATGC GAATCTCATC TATTTTCTCT	120
TTTTTTTGCA ATTTAGCTGA TTTTCTTTT TTACCATTTA CAGTCACGCG CCCAGCCTTG	180
AGCAAGTTTT TGACCTCAGT CCGACTTCCC ACCGCACAGG CAACTAAAAA TTTATCTAAT	240
CTCATAGAAC TATTATATCA TATCAAAAGG AGGCTAGTAC AATGACCAAC CTCCTTTTCG	300
TTTCATACTC TTCAAAAATC TCTTCAAACC GCGTCAACGT CGCCTTGCCG TATATATGTT	360
ACTGACTTCG TCAGTTCTAT CTGCAACCTC AAAACAGTGT TTTGAGCTGA CTCGTCAGT	420
TCTATCTGCA ACCTCAAAGC AGTGCTTTGA GCATCCTGCG GCTAGTTTCC kAGTktGCTC	480
TTTGATTWC ATTGAGTATC AGATTTAGGA AATTAAC TTCGkCTCCA AAAAAKAGCT	540
AAAACAATCA AGGCTCCTAA AATCGCTGGG AT	572

(2) INFORMATION FOR SEQ ID NO: 389:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 505 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 389:

CAACAAGTGC CTCGGCTTCA GCATGCACAA GTGCTTCAGC TTCAGCATGT ACCTGAGCGT	60
CTGAATCAGC ATCAACGTGT GCGTCCGCTT CAGCATGTAC TGCTGCCTCA GCATCAGCGT	120
CAACAwGTGC TTCGGCTTCA GCGTCAACGA GTGCGTCTGA GTCAGCATCA ACGAGTACGT	180
CAGCCTCAGC AAGCACATCA GCTTCTGAAT CTGCATCAAC CAGTGCGTCA GCCTCAGCAT	240
CGACAAGCGC CTCAGCTTCA GCAAGTACCA GTGCGTCAGC CTCAGCAAGT ACCAGTGCTT	300
CAGCCTCAGC GTCGACAAGT GCGTCGCGCT CAACCAGTGC ATCTGAATCG GCATCAACCA	360
GTGCGTCAGC CTCAGCAAGT ACTAGCGCCT CAGCCTCAGC ATCAACGAGT GCGTCCGCTT	420

1397

CAGCAAGTAC TAGTGCATCA GCATCAGCAT CAACGAGTGC ATCGGCTTCA GCAAGTACCA 480
GCGCCTCAGC TTCAGCAAGC ACCGG 505

(2) INFORMATION FOR SEQ ID NO: 390:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 447 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 390:

GCTAAGACTA CCTCATTAGG GGCATAGGCT GCTAAAATAA CTGCAGCTGT GGTTAATGAC 60
AATACTGTAC TTTTTCAT TTTAATCCT TACATATTTA TATAACTTCC AATAGATAAT 120
AAACTTTAAC TTTGCTAGCC TTTGTTATAA AAAGTTTAC TAAGTATTAT CTAGGAAATA 180
GAGTAGTACA TTTATATATA ATTGTTATCT CTCTATAAAA ACAGTATATC ATTTAAAAA 240
ATTTAAGTCA AAAAAATTAA CATTAGTTAA TTTATTTTTT AGCACACATT AAAAAATAAG 300
ATTAGTACTC AATGAAAATC AAAGAGCAAA CTAGGAAACT AGCCGCAGAT TGCTCAAAAC 360
AGTGTTTTGA GGTGTAGAT GGAATGACGT AGTCAGCTCA AAACACTGTT TTGAAGTTGT 420
GGATAGAACT GACGAAGTCG GTACCGA 447

(2) INFORMATION FOR SEQ ID NO: 391:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 572 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 391:

AGCACTTGTC GTTGAATTCT ACAACAAAAT GTTGAATAT TTTATTGAAT AAGATAGGCC 60
TTGATATTAA GCACTTTGGG ACGTTCCTCC TTAGTGCTTT TTTGATTCT CTAGTATCC 120
AGCTATAATC GTTGAGACAT AACTAGACCG ATATAGTCCA AAGTGATATA GTAAAATGAA 180
CCAAAAATAG TACACAATGT GGTATAATCC TTTTATGGCA TATTCAATAG ATTTTCGTAA 240
AAAAGTTCTC TCTTATTGTG AGCGAACAGG TAGTATAACA GAAGCATCAC ACGTTTCCCA 300
AATCTCACGT AATACCATTT ATGGCTGGTT AAAGCTAAAA GAGAAAACAG GAGAGCTAAA 360
CCACCAAGTA TAGTGTATTG AATCTATAAC AGTACACCTT GGCTGCTAAA ATATTCTAT 420

1398
AAATTAATTT GACTTTCCTG ATAGAGATGT TCACATCTTA TTCAAATA CTATATAAGT 480
TCTATAATCT CTTTATAAGA TTGCCCATC AGACAAAATA GAACGATTG AAGGCGTTTA 540
TGATATTTAG CTGTACGAGA GTCTTTTAAA AG 572

MISSING UPON TIME OF PUBLICATION

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person approved by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PUT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the International publication of the application.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapse, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever two dates occurs earlier.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegians Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

ICELAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected in the art.

What Is Claimed Is:

- 25 1. Computer readable medium having recorded thereon the nucleotide sequence depicted in SEQ ID NOS:1-391, a representative fragment thereof or a nucleotide sequence at least 95% identical to a nucleotide sequence depicted in SEQ ID NOS:1-391.
- 30 2. Computer readable medium having recorded thereon any one of the fragments of SEQ ID NOS:1-391 depicted in Tables 2 and 3 or a degenerate variant thereof.
- 35 3. The computer readable medium of claim 1, wherein said medium is selected from the group consisting of a floppy disc, a hard disc, random access memory (RAM), read only memory (ROM), and CD-ROM.
- 40 4. The computer readable medium of claim 3, wherein said medium is selected from the group consisting of a floppy disc, a hard disc, random access memory (RAM), read only memory (ROM), and CD-ROM.
5. A computer-based system for identifying fragments of the *Streptococcus pneumoniae* genome of commercial importance comprising the following elements:
- 45 a) a data storage means comprising the nucleotide sequence of SEQ ID NOS:1-391, a representative fragment thereof, or a nucleotide sequence at least 95% identical to a nucleotide sequence of SEQ ID NOS:1-391;
- b) search means for comparing a target sequence to the nucleotide sequence of the data storage means of step (a) to identify homologous sequence(s), and
- c) retrieval means for obtaining said homologous sequence(s) of step (b).
- 50 6. A method for identifying commercially important nucleic acid fragments of the *Streptococcus pneumoniae* genome comprising the step of comparing a database comprising the nucleotide sequences depicted in SEQ ID NOS:1-391, a representative fragment thereof, or a nucleotide sequence at least 95% identical to a nucleotide sequence of SEQ ID NOS:1-391 with a target sequence to obtain a nucleic acid molecule comprised of a complementary nucleotide sequence to said target sequence, wherein said target sequence is not randomly selected.
- 55

60 7. A method for identifying an expression modulating fragment of
Streptococcus pneumoniae genome comprising the step of comparing a database
comprising the nucleotide sequences depicted in SEQ ID NOS:1-391, a
representative fragment thereof, or a nucleotide sequence at least 95% identical to
the nucleotide sequence of SEQ ID NOS:1-391 with a target sequence to obtain a
nucleic acid molecule comprised of a complementary nucleotide sequence to said
65 target sequence, wherein said target sequence comprises sequences known to
regulate gene expression.

70 8. An isolated protein-encoding nucleic acid fragment of the *Streptococcus*
pneumoniae genome, wherein said fragment consists of the nucleotide sequence of
any one of the fragments of SEQ ID NOS:1-391 depicted in Tables 2 and 3, or a
degenerate variant thereof.

75 9. A vector comprising any one of the fragments of the *Streptococcus*
pneumoniae genome SEQ ID NOS:1-391 depicted in Tables 2 and 3 or a
degenerate variant thereof.

80 10. An isolated fragment of the *Streptococcus pneumoniae* genome,
wherein said fragment modulates the expression of an operably linked open reading
frame, wherein said fragment consists of the nucleotide sequence from about 10 to
200 bases in length which is 5' to any one of the open reading frames depicted in
Tables 2 and 3 or a degenerate variant thereof.

85 11. A vector comprising any one of the fragments of the *Streptococcus*
pneumoniae genome of claim 8.

12. An organism which has been altered to contain any one of the
fragments of the *Streptococcus pneumoniae* genome of claim 8.

90 13. An organism which has been altered to contain any one of the
fragments of the *Streptococcus pneumoniae* genome of claim 10.

95 14. A method for regulating the expression of a nucleic acid molecule comprising the step of covalently attaching to said nucleic acid molecule a nucleic acid molecule consisting of the nucleotide sequence from about 10 to 100 bases 5' to any one of the fragments of the *Streptococcus pneumoniae* genome depicted in SEQ ID NOS:1-391 and Tables 2 and 3 or a degenerate variant thereof.

100 15. An isolated nucleic acid molecule encoding a homolog of any of the fragments of the *Streptococcus pneumoniae* genome of SEQ ID NOS:1-391 and Tables 2 and 3, wherein said nucleic acid molecule is produced by a process comprising steps of:

- 105 a) screening a genomic DNA library using as a probe a target sequence defined by any of SEQ ID NOS:1-391 and Tables 2 and 3, including fragments thereof;
- b) identifying members of said library which contain sequences that hybridize to said target sequence; and
- c) isolating the nucleic acid molecules from said members identified in step (b).

110 16. An isolated DNA molecule encoding a homolog of any one of the fragments of the *Streptococcus pneumoniae* genome of SEQ ID NOS:1-391 and Tables 2 and 3, wherein said nucleic acid molecule is produced a process comprising steps of:

- 115 a) isolating mRNA, DNA, or cDNA produced from an organism;
- b) amplifying nucleic acid molecules whose nucleotide sequence is homologous to amplification primers derived from said fragment of said *Streptococcus pneumoniae* genome to prime said amplification;
- c) isolating said amplified sequences produced in step (b).

120 17. An isolated polypeptide encoded by any of the fragments of the *Streptococcus pneumoniae* genome of SEQ ID NOS:1-391 and depicted in Table 2 and 3 or by a degenerate variant of said fragments.

125 18. An isolated polynucleotide molecule encoding any one of the polypeptides of claim 17.

19. An antibody which selectively binds to any one of the polypeptides of claim 17.

130

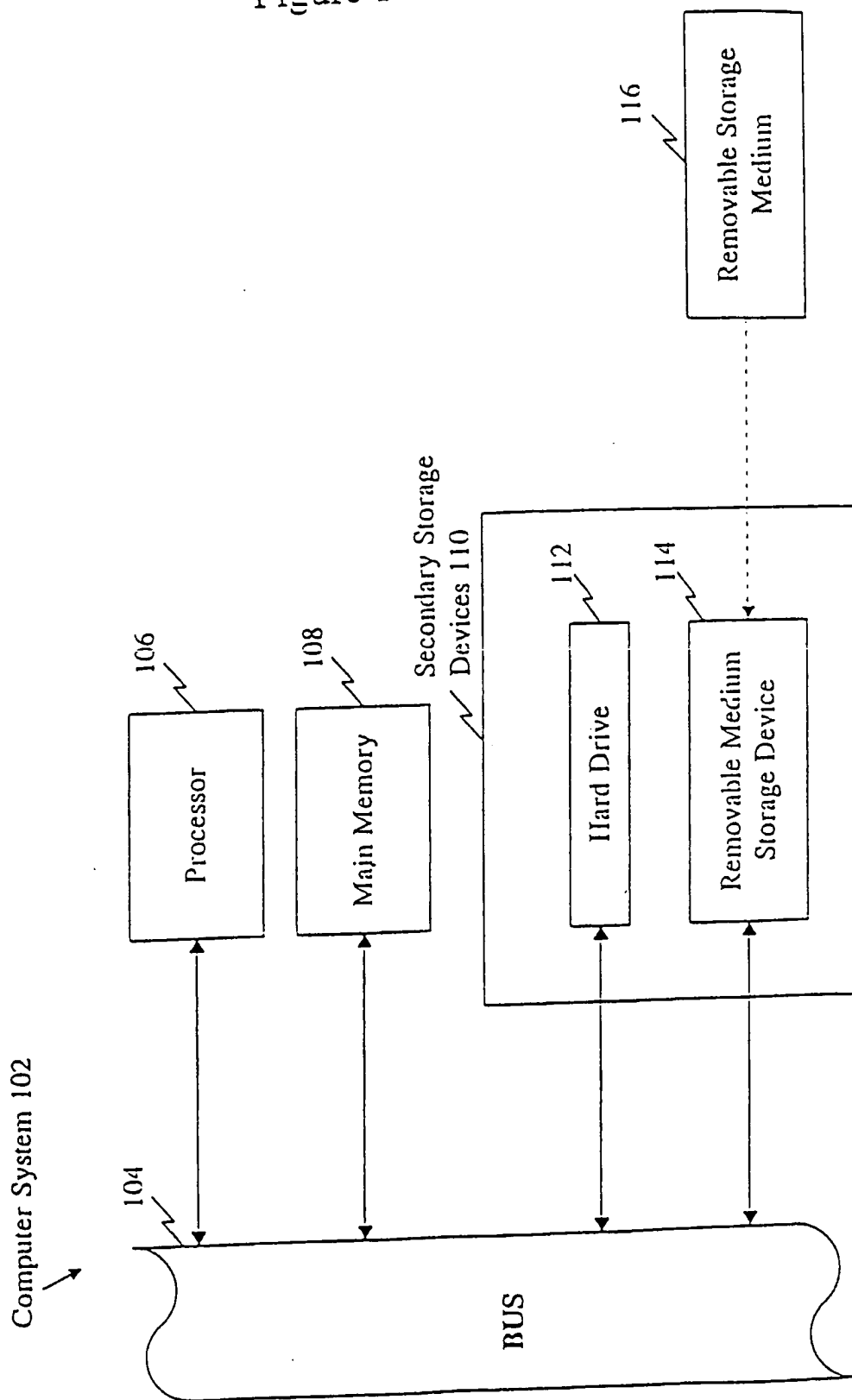
20. A method for producing a polypeptide in a host cell comprising the steps of:

135

a) incubating a host containing a heterologous nucleic acid molecule whose nucleotide sequence consists of any one of the fragments of the *Streptococcus pneumoniae* genome of SEQ ID NOS:1-391 and depicted in Tables 2 and 3, under conditions where said heterologous nucleic acid molecule is expressed to produce said protein, and

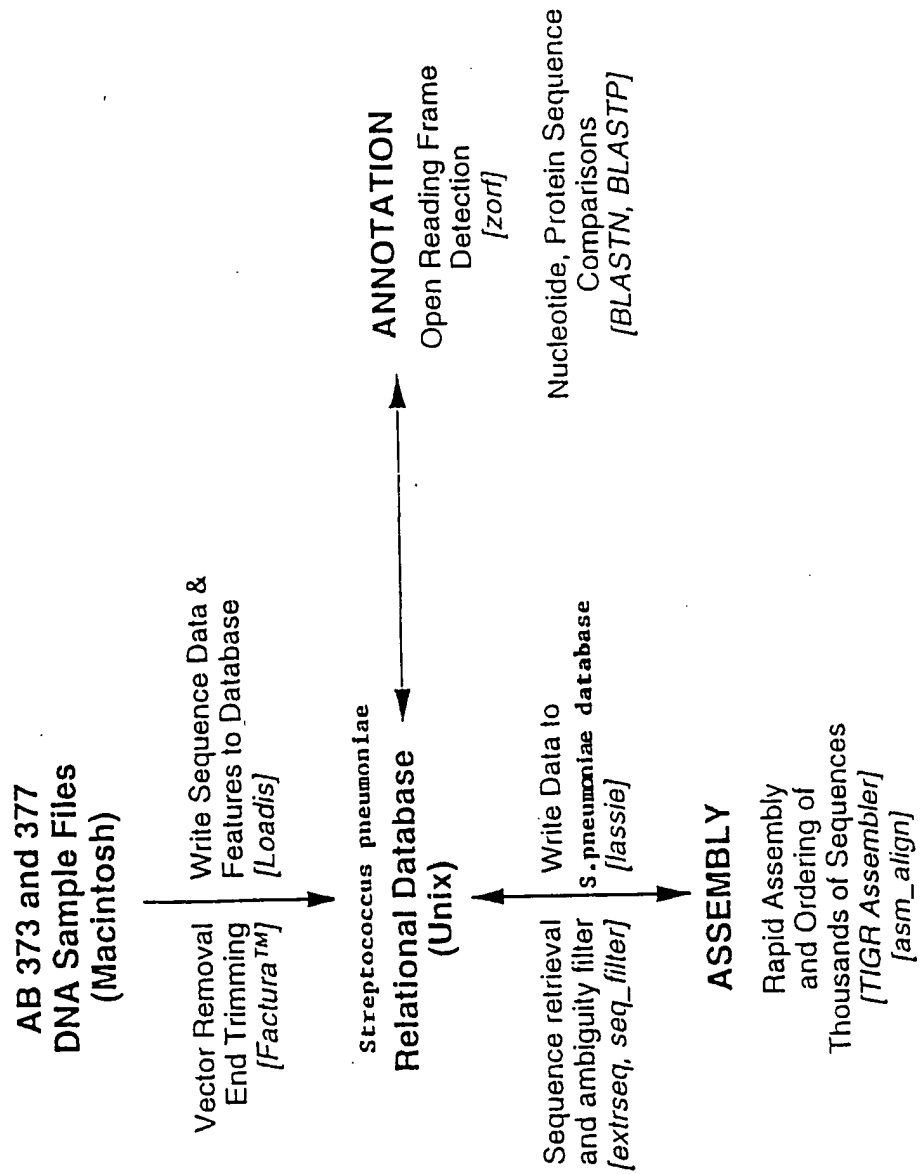
b) isolating said protein.

Figure 1



2/2

Figure 2



PCT

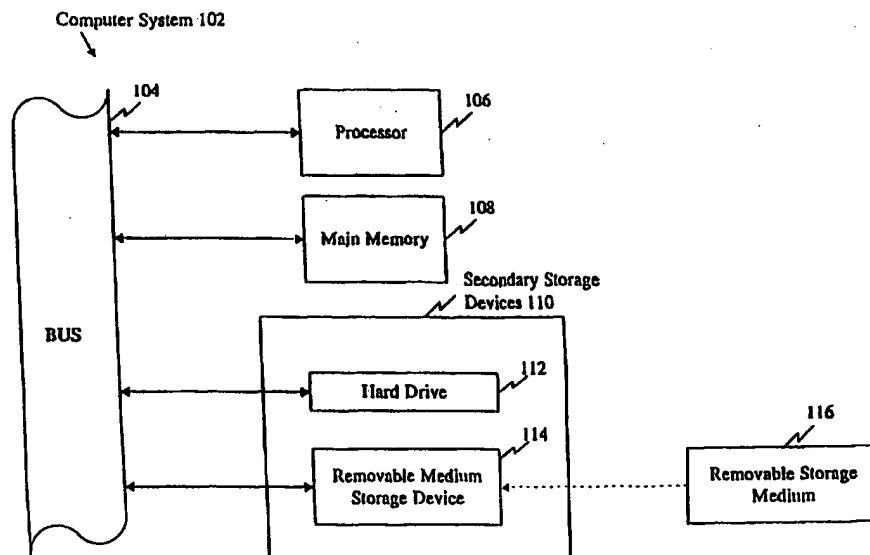
WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/31, C07K 14/315, 16/12, C12Q 1/68		(11) International Publication Number: WO 98/18931
A3		(43) International Publication Date: 7 May 1998 (07.05.98)
(21) International Application Number: PCT/US97/19588		(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).
(22) International Filing Date: 30 October 1997 (30.10.97)		
(30) Priority Data: 60/029,960 31 October 1996 (31.10.96) US		
(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).		
(72) Inventors; and (75) Inventors/Applicants (for US only): KUNSCH, Charles, A. [US/US]; 2398B Dunwoody Crossing, Atlanta, GA 30338 (US). CHOI, Gil, H. [KR/US]; 11429 Potomac Oaks Drive, Rockville, MD 20850 (US). DILLON, Patrick, J. [US/US]; 1055 Snipe Court, Carlsbad, CA 92009 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). BARASH, Steven, C. [US/US]; 582 College Parkway #303, Rockville, MD 20850 (US). FANNON, Michael [US/US]; 13501 Rippling Brook Drive, Silver Spring, MD 20850 (US). DOUGHERTY, Brian, A. [US/US]; 708 Meadow Field Court, Mount Airy, MD 21771 (US).		
		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).
		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.
		(88) Date of publication of the international search report: 20 August 1998 (20.08.98)

(54) Title: *STREPTOCOCCUS PNEUMONIAE* POLYNUCLEOTIDES AND SEQUENCES



(57) Abstract

The present invention provides polynucleotide sequences of the genome of *Streptococcus pneumoniae*, polypeptide sequences encoded by the polynucleotide sequences, corresponding polynucleotides and polypeptides, vectors and hosts comprising the polynucleotides, and assays and other uses thereof. The present invention further provides polynucleotide and polypeptide sequence information stored on computer readable media, and computer-based systems and methods which facilitate its use.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 97/19588

A. CLASSIFICATION F. SUBJECT MATTER

IPC 6 C12N15/31 C07K14/315 C07K16/12 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 96 33276 A (HUMAN GENOME SCIENCES INC ;UNIV JOHNS HOPKINS (US)) 24 October 1996 see claims 1-7	1-7
A	--- ALTSCHUL S F ET AL: "BASIL LOCAL ALIGNMENT SEARCH TOOL" JOURNAL OF MOLECULAR BIOLOGY, vol. 215, 1990, pages 403-410, XP000604562 cited in the application see the whole document --- -/--	1-7

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

27 March 1998

Date of mailing of the international search report

08.07.98

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

HORNIG H.

INTERNATIONAL SEARCH REPORT

Int. l.ional Application No

PCT/US 97/19588

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	W.R. PEARSON AND D.J. LIPMAN: "Improved tools for biological sequence comparison" PROC. NATL. ACAD. SCI., vol. 85, April 1988, NATL. ACAD. SCI., WASHINGTON, DC, US;, pages 2444-2448, XP002060460 cited in the application see the whole document ---	1-7
A	WO 95 06732 A (UNIV ROCKEFELLER ;MASURE H ROBERT (US); PEARCE BARBARA J (US); TUO) 9 March 1995 see the whole document ---	1-7
A	WO 95 31548 A (UAB RESEARCH FOUNDATION ;YOTHER JANET (US); DILLARD JOSEPH P (US)) 23 November 1995 see the whole document ---	1-7
A	WO 95 14712 A (RES CORP TECHNOLOGIES INC) 1 June 1995 see the whole document ---	1-7
A	WO 96 05859 A (AMERICAN CYANAMID CO) 29 February 1996 see the whole document ---	1-7
A	WO 93 10238 A (US HEALTH) 27 May 1993 see the whole document ---	1-7
A	EP 0 687 688 A (UNIV OVIEDO ;UNIV LEICESTER (GB)) 20 December 1995 see the whole document ---	1-7
A	EP 0 622 081 A (UAB RESEARCH FOUNDATION) 2 November 1994 see the whole document -----	1-7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 97/19588

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 1-4 could be, at least partially be considered as a mere presentation of information Rule 39.1(v), and claims 5-7 at least partially as a computer program, Rule 39.1(vi)PCT, the search has been carried out as far as possible in our systematic documentation.
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see continuation-sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-7

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-7

Computer readable medium having recorded thereon the nucleotide sequence depicted in SEQ ID nos. 1-391, a representative fragment thereof or a nucleotide sequence at least 95% identical to a nucleotide sequence depicted in SEQ ID nos. 1-391; a computer-based system for identifying fragments of the *Streptococcus pneumoniae* genome of commercial importance comprising: a) a data storage means comprising said nucleotide sequence(s); b) search means for comparing a target sequence to the nucleotide sequence of the data storage means of step (a) to identify homologous sequence(s), and c) retrieval means for obtaining said homologous sequence(s) of step (b); a method for identifying commercially important nucleic acid fragments of the *Streptococcus pneumoniae* genome comprising the step of comparing a database comprising said nucleotide sequence(s) with a target sequence to obtain a nucleic acid molecule comprised of a complementary nucleotide sequence to said target sequence, wherein said target sequence is not randomly selected; a method for identifying an expression modulating fragments of the *Streptococcus pneumoniae* genome comprising the step of comparing a database comprising said nucleotide sequence(s) with a target sequence to obtain a nucleic acid molecule comprised of a complementary nucleotide sequence to said target sequence, wherein said target sequence comprises sequences known to regulate gene expression;

2. Claims: (8-20) partially

An isolated protein-encoded nucleic acid fragment of the *Streptococcus pneumoniae* genome, wherein said fragment consists of the nucleotide sequence of the fragment of SEQ ID no.1 depicted in Tables 2 and 3, or a degenerate variant thereof; a vector comprising the fragment of the *Streptococcus pneumoniae* genome SEQ ID no.1; an isolated fragment of the *Streptococcus pneumoniae* genome, wherein said fragment modulates the expression of an operably linked open reading frame, wherein said fragment consists of the nucleotide sequence from about 10 to 200 bases in length which is 5' to any one of the open reading frame of SEQ ID no.1 depicted in Tables 2 and 3 or a degenerate variant thereof; a method for regulating the expression of a nucleic acid molecule comprising the step of covalently attaching to said nucleic acid molecule a nucleic acid molecule consisting of the nucleotide sequence from about 10 to 100 bases 5' to any one of the open reading frame of SEQ ID no.1 and Tables 2 and 3 or a degenerate variant thereof; an isolated nucleic acid molecule encoding a homolog of SEQ ID no.1; an isolated polypeptide encoded by SEQ ID no.1 and depicted in Table 2 and 3; an antibody which selectively binds to any one of said polypeptides, a method for producing a polypeptide in a host cell comprising a) incubating a host containing a heterologous nucleic acid

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

molecule whose nucleotide sequence consists of SEQ ID no.1 and depicted in Table 2 and 3, under conditions where said heterologous nucleic acid molecule is expressed to produce said protein, and b) isolating said protein;

3-392. Claims:(8-20) partially

Idem as subject 2 but limited to each of the sequences of SEQ ID no. 2 to 391;

For the sake of conciseness, the second subject matter is explicitly defined, the other subject matters are defined by analogy hereto.

INTERNATIONAL SEARCH REPORT

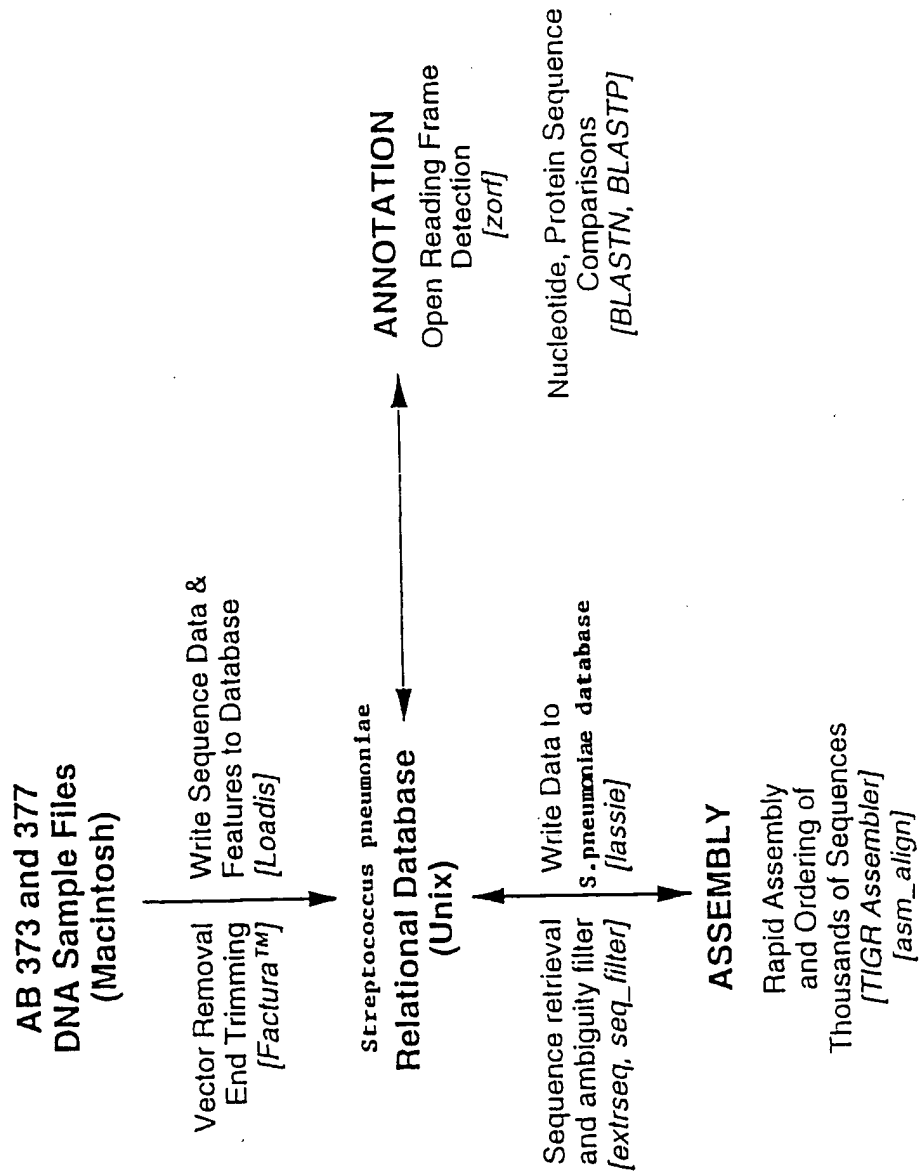
Information on patent family members

International Application No

PCT/US 97/19588

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9633276 A	24-10-96	AU 5552396 A EP 0821737 A	07-11-96 04-02-98
WO 9506732 A	09-03-95	AU 7680994 A CA 2170726 A EP 0721506 A FI 960977 A JP 9504686 T NO 960839 A	22-03-95 09-03-95 17-07-96 30-04-96 13-05-97 19-04-96
WO 9531548 A	23-11-95	AU 2638595 A EP 0804582 A	05-12-95 05-11-97
WO 9514712 A	01-06-95	US 5474905 A	12-12-95
WO 9605859 A	29-02-96	US 5565204 A AU 3363695 A CA 2198251 A EP 0778781 A JP 10504717 T	15-10-96 14-03-96 29-02-96 18-06-97 12-05-98
WO 9310238 A	27-05-93	AU 3065892 A	15-06-93
EP 0687688 A	20-12-95	ES 2075803 A ES 2088820 A WO 9516711 A	01-10-95 16-09-96 22-06-95
EP 0622081 A	02-11-94	AU 682018 B AU 5769694 A CA 2116261 A FI 941695 A JP 7126291 A NO 941420 A US 5679768 A ZA 9401584 A	18-09-97 27-10-94 21-10-94 21-10-94 16-05-95 21-10-94 21-10-97 12-10-94

Figure 2

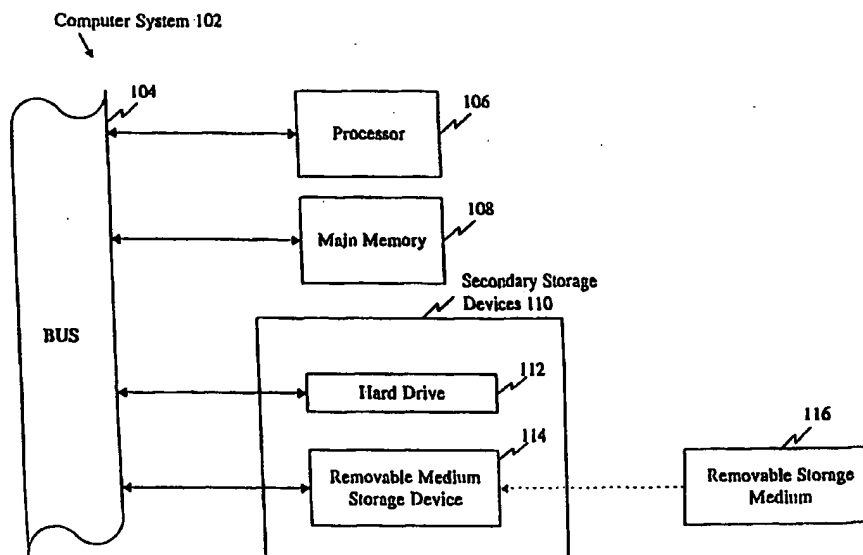


PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/31, C07K 14/315, 16/12, C12Q 1/68		(11) International Publication Number: WO 98/18931
A3		(43) International Publication Date: 7 May 1998 (07.05.98)
(21) International Application Number: PCT/US97/19588		(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).
(22) International Filing Date: 30 October 1997 (30.10.97)		
(30) Priority Data: 60/029,960 31 October 1996 (31.10.96) US		
(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).		
(72) Inventors; and (75) Inventors/Applicants (for US only): KUNSCH, Charles, A. [US/US]; 2398B Dunwoody Crossing, Atlanta, GA 30338 (US). CHOI, Gil, H. [KR/US]; 11429 Potomac Oaks Drive, Rockville, MD 20850 (US). DILLON, Patrick, J. [US/US]; 1055 Snipe Court, Carlsbad, CA 92009 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). BARASH, Steven, C. [US/US]; 582 College Parkway #303, Rockville, MD 20850 (US). FANNON, Michael [US/US]; 13501 Rippling Brook Drive, Silver Spring, MD 20850 (US). DOUGHERTY, Brian, A. [US/US]; 708 Meadow Field Court, Mount Airy, MD 21771 (US).		
		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).
		Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
		(88) Date of publication of the international search report: 20 August 1998 (20.08.98)

(54) Title: **STREPTOCOCCUS PNEUMONIAE POLYNUCLEOTIDES AND SEQUENCES**

(57) Abstract

The present invention provides polynucleotide sequences of the genome of *Streptococcus pneumoniae*, polypeptide sequences encoded by the polynucleotide sequences, corresponding polynucleotides and polypeptides, vectors and hosts comprising the polynucleotides, and assays and other uses thereof. The present invention further provides polynucleotide and polypeptide sequence information stored on computer readable media, and computer-based systems and methods which facilitate its use.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 97/19588

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/31 C07K14/315 C07K16/12 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 96 33276 A (HUMAN GENOME SCIENCES INC ;UNIV JOHNS HOPKINS (US)) 24 October 1996 see claims 1-7	1-7
A	--- ALTSCHUL S F ET AL: "BASIL LOCAL ALIGNMENT SEARCH TOOL" JOURNAL OF MOLECULAR BIOLOGY, vol. 215, 1990, pages 403-410, XP000604562 cited in the application see the whole document --- -/--	1-7



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

27 March 1998

Date of mailing of the international search report

08.07.98

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

HORNIG H.

INTERNATIONAL SEARCH REPORT

Int'l. Application No

PCT/US 97/19588

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	W.R. PEARSON AND D.J. LIPMAN: "Improved tools for biological sequence comparison" PROC. NATL. ACAD. SCI., vol. 85, April 1988, NATL. ACAD. SCI., WASHINGTON, DC, US;, pages 2444-2448, XP002060460 cited in the application see the whole document ---	1-7
A	WO 95 06732 A (UNIV ROCKEFELLER ;MASURE H ROBERT (US); PEARCE BARBARA J (US); TUO) 9 March 1995 see the whole document ---	1-7
A	WO 95 31548 A (UAB RESEARCH FOUNDATION ;YOTHER JANET (US); DILLARD JOSEPH P (US)) 23 November 1995 see the whole document ---	1-7
A	WO 95 14712 A (RES CORP TECHNOLOGIES INC) 1 June 1995 see the whole document ---	1-7
A	WO 96 05859 A (AMERICAN CYANAMID CO) 29 February 1996 see the whole document ---	1-7
A	WO 93 10238 A (US HEALTH) 27 May 1993 see the whole document ---	1-7
A	EP 0 687 688 A (UNIV OVIEDO ;UNIV LEICESTER (GB)) 20 December 1995 see the whole document ---	1-7
A	EP 0 622 081 A (UAB RESEARCH FOUNDATION) 2 November 1994 see the whole document -----	1-7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 97/ 19588

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 1-4 could be, at least partially be considered as a mere presentation of information Rule 39.1(v), and claims 5-7 at least partially as a computer program, Rule 39.1(vi)PCT, the search has been carried out as far as possible in our systematic documentation.
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see continuation-sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-7

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-7

Computer readable medium having recorded thereon the nucleotide sequence depicted in SEQ ID nos. 1-391, a representative fragment thereof or a nucleotide sequence at least 95% identical to a nucleotide sequence depicted in SEQ ID nos. 1-391; a computer-based system for identifying fragments of the *Streptococcus pneumoniae* genome of commercial importance comprising: a) a data storage means comprising said nucleotide sequence(s); b) search means for comparing a target sequence to the nucleotide sequence of the data storage means of step (a) to identify homologous sequence(s), and c) retrieval means for obtaining said homologous sequence(s) of step (b); a method for identifying commercially important nucleic acid fragments of the *Streptococcus pneumoniae* genome comprising the step of comparing a database comprising said nucleotide sequence(s) with a target sequence to obtain a nucleic acid molecule comprised of a complementary nucleotide sequence to said target sequence, wherein said target sequence is not randomly selected; a method for identifying an expression modulating fragments of the *Streptococcus pneumoniae* genome comprising the step of comparing a database comprising said nucleotide sequence(s) with a target sequence to obtain a nucleic acid molecule comprised of a complementary nucleotide sequence to said target sequence, wherein said target sequence comprises sequences known to regulate gene expression;

2. Claims: (8-20) partially

An isolated protein-encoded nucleic acid fragment of the *Streptococcus pneumoniae* genome, wherein said fragment consists of the nucleotide sequence of the fragment of SEQ ID no.1 depicted in Tables 2 and 3, or a degenerate variant thereof; a vector comprising the fragment of the *Streptococcus pneumoniae* genome SEQ ID no.1; an isolated fragment of the *Streptococcus pneumoniae* genome, wherein said fragment modulates the expression of an operably linked open reading frame, wherein said fragment consists of the nucleotide sequence from about 10 to 200 bases in length which is 5' to any one of the open reading frame of SEQ ID no.1 depicted in Tables 2 and 3 or a degenerate variant thereof; a method for regulating the expression of a nucleic acid molecule comprising the step of covalently attaching to said nucleic acid molecule a nucleic acid molecule consisting of the nucleotide sequence from about 10 to 100 bases 5' to any one of the open reading frame of SEQ ID no.1 and Tables 2 and 3 or a degenerate variant thereof; an isolated nucleic acid molecule encoding a homolog of SEQ ID no.1; an isolated polypeptide encoded by SEQ ID no.1 and depicted in Table 2 and 3; an antibody which selectively binds to any one of said polypeptides, a method for producing a polypeptide in a host cell comprising a) incubating a host containing a heterologous nucleic acid

FURTHER INFORMATION CONTINUED FROM PCT/SA/ 210

molecule whose nucleotide sequence consists of SEQ ID no.1 and depicted in Table 2 and 3, under conditions where said heterologous nucleic acid molecule is expressed to produce said protein, and b) isolating said protein;

3-392. Claims: (8-20) partially

Idem as subject 2 but limited to each of the sequences of SEQ ID no. 2 to 391;

For the sake of conciseness, the second subject matter is explicitly defined, the other subject matters are defined by analogy hereto.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-7

Computer readable medium having recorded thereon the nucleotide sequence depicted in SEQ ID nos. 1-391, a representative fragment thereof or a nucleotide sequence at least 95% identical to a nucleotide sequence depicted in SEQ ID nos. 1-391; a computer-based system for identifying fragments of the *Streptococcus pneumoniae* genome of commercial importance comprising: a) a data storage means comprising said nucleotide sequence(s); b) search means for comparing a target sequence to the nucleotide sequence of the data storage means of step (a) to identify homologous sequence(s), and c) retrieval means for obtaining said homologous sequence(s) of step (b); a method for identifying commercially important nucleic acid fragments of the *Streptococcus pneumoniae* genome comprising the step of comparing a database comprising said nucleotide sequence(s) with a target sequence to obtain a nucleic acid molecule comprised of a complementary nucleotide sequence to said target sequence, wherein said target sequence is not randomly selected; a method for identifying an expression modulating fragments of the *Streptococcus pneumoniae* genome comprising the step of comparing a database comprising said nucleotide sequence(s) with a target sequence to obtain a nucleic acid molecule comprised of a complementary nucleotide sequence to said target sequence, wherein said target sequence comprises sequences known to regulate gene expression;

2. Claims: (8-20) partially

An isolated protein-encoded nucleic acid fragment of the *Streptococcus pneumoniae* genome, wherein said fragment consists of the nucleotide sequence of the fragment of SEQ ID no.1 depicted in Tables 2 and 3, or a degenerate variant thereof; a vector comprising the fragment of the *Streptococcus pneumoniae* genome SEQ ID no.1; an isolated fragment of the *Streptococcus pneumoniae* genome, wherein said fragment modulates the expression of an operably linked open reading frame, wherein said fragment consists of the nucleotide sequence from about 10 to 200 bases in length which is 5' to any one of the open reading frame of SEQ ID no.1 depicted in Tables 2 and 3 or a degenerate variant thereof; a method for regulating the expression of a nucleic acid molecule comprising the step of covalently attaching to said nucleic acid molecule a nucleic acid molecule consisting of the nucleotide sequence from about 10 to 100 bases 5' to any one of the open reading frame of SEQ ID no.1 and Tables 2 and 3 or a degenerate variant thereof; an isolated nucleic acid molecule encoding a homolog of SEQ ID no.1; an isolated polypeptide encoded by SEQ ID no.1 and depicted in Table 2 and 3; an antibody which selectively binds to any one of said polypeptides, a method for producing a polypeptide in a host cell comprising a) incubating a host containing a heterologous nucleic acid